

GenCore version: 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2004, 03:25:04 ; Search time 925 Seconds
(without alignments)
11198.630 Million cell updates/sec

Title: US-09-765-272A-65

Perfect score: 2290

Sequence: 1 TTGCTCTATGAAGTGGTC.....TACTTGAAAAATCCAGGAT 2290

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
1	2285	99.8	2290	9	US-09-765-272-65
2	2244.8	98.0	2481	10	US-09-769-787-206
3	2220.2	97.0	2531	13	US-10-412-862-11
4	2220.2	97.0	2531	16	US-10-412-862-11
5	2220.2	97.0	2531	16	US-10-412-862-11
6	2022.2	98.3	2639	10	US-09-884-465A-5
7	1999.6	87.3	2531	13	US-10-412-862-5
8	1999.6	87.3	2531	16	US-10-412-862-5
9	1999.6	87.3	2531	16	US-10-387-783-5
10	1999.2	87.3	8195	13	US-10-158-844-94
11	1856.8	81.1	2523	10	US-09-884-465A-3
12	1856.8	81.1	2647	10	US-09-884-465A-4
13	1376.6	60.1	2359	13	US-10-158-844-243
14	990.6	43.3	2451	13	US-10-412-862-9

15	990.6	43.3	2451	16	US-10-412-850-9	Sequence 9, Appli
16	990.6	43.3	2451	16	US-10-387-783-9	Sequence 9, Appli
17	987.6	43.1	2389	9	US-09-765-272-55	Sequence 55, Appli
18	794.8	34.7	973	13	US-10-158-844-355	Sequence 355, App
19	677	29.6	1684	13	US-10-158-844-258	Sequence 258, App
20	401	17.5	1342	9	US-09-765-272-181	Sequence 181, App
21	401	17.5	1455	10	US-09-769-787-246	Sequence 246, App
22	401	17.5	1455	10	US-09-769-787-246	Sequence 23, Appli
23	401	17.5	1455	13	US-10-412-862-7	Sequence 7, Appli
24	401	17.5	1455	16	US-10-412-850-7	Sequence 7, Appli
25	401	17.5	1455	16	US-10-387-783-7	Sequence 7, Appli
26	401	17.5	3120	10	US-09-884-465A-1	Sequence 1, Appli
27	401	17.5	5048	10	US-09-884-465A-2	Sequence 2, Appli
28	401	17.5	6867	13	US-10-158-844-192	Sequence 192, App
29	370	16.2	819	10	US-09-884-465A-257	Sequence 257, App
30	370	16.2	819	16	US-10-324-143-10	Sequence 10, Appli
31	368.4	16.1	816	16	US-10-324-143-4	Sequence 4, Appli
32	368.4	16.1	816	16	US-10-324-143-5	Sequence 5, Appli
33	363.6	15.9	2721	16	US-10-324-143-6	Sequence 6, Appli
34	346.6	15.1	2528	10	US-09-884-465A-9	Sequence 9, Appli
35	272.4	11.9	841	9	US-09-452-599-34	Sequence 34, Appli
36	272.4	11.9	841	15	US-10-121-120-34	Sequence 34, Appli
37	225	9.8	1146	10	US-09-769-736-23	Sequence 23, Appli
38	225	9.8	2469	10	US-09-769-736-17	Sequence 17, Appli
39	223.4	9.8	5215	10	US-09-252-088-13	Sequence 13, Appli
40	223.4	9.8	5215	16	US-10-340-792-13	Sequence 13, Appli
41	117	5.1	1455	10	US-09-769-736-71	Sequence 71, Appli
42	58.2	2.5	31124	13	US-10-087-192-463	Sequence 463, App
43	54.8	2.4	45980	10	US-09-957-956-6	Sequence 6, Appli
44	54	2.4	2232	15	US-10-087-464-45	Sequence 45, Appli
45	52.6	2.3	635	13	US-10-027-632-269927	Sequence 269927,
46	52.6	2.3	635	16	US-10-027-632-269927	Sequence 269927,
47	52.2	2.3	96596	12	US-10-052-483-70	Sequence 70, Appli
48	51.2	2.2	305	9	US-09-864-761-19262	Sequence 19262, A
49	51.2	2.2	496	9	US-09-864-761-2534	Sequence 2534, Ap
50	50.6	2.2	5928	13	US-10-221-714A-492	Sequence 492, App
51	50.6	2.2	5928	15	US-10-311-455-2212	Sequence 2212, Ap
52	50	2.2	1010	15	US-10-029-386-24901	Sequence 24901, A
53	49.8	2.2	122859	13	US-10-087-193-37	Sequence 37, Appli
54	49.6	2.2	678	9	US-09-864-761-20595	Sequence 20595, A
55	49.4	2.2	678	13	US-10-027-632-141716	Sequence 141716,
56	49.4	2.2	678	16	US-10-027-632-141716	Sequence 141716,
57	49.2	2.1	438	9	US-09-864-761-4988	Sequence 4988, Ap
58	49.2	2.1	2499	9	US-09-842-552-92	Sequence 92, Appli
59	49.2	2.1	2550	15	US-10-032-585-6026	Sequence 6026, Ap
60	48.8	2.1	408	9	US-09-360-352-1221	Sequence 1221, Ap
61	48.8	2.1	524	13	US-10-027-632-50439	Sequence 50439, A
62	48.8	2.1	524	16	US-10-027-632-50439	Sequence 50439, A
63	48.4	2.1	37285	13	US-10-087-192-49	Sequence 49, Appli
64	48.4	2.1	374849	13	US-10-087-192-1627	Sequence 1627, Ap
65	48.2	2.1	2276	13	US-10-027-632-110545	Sequence 110545,
66	48.2	2.1	2276	16	US-10-027-632-110545	Sequence 110545,
67	48.2	2.1	14798	15	US-10-311-455-1006	Sequence 1006, Ap
68	48	2.1	3912	13	US-10-282-122A-9928	Sequence 9928, Ap
69	48	2.1	96597	12	US-10-052-482-103	Sequence 103, App
70	47.8	2.1	315	9	US-09-864-761-21723	Sequence 21723, A
71	47.8	2.1	18846	9	US-09-815-242-8898	Sequence 8898, Ap
72	47.8	2.1	18846	8	US-08-781-986A-59	Sequence 59, Appli
73	47.8	2.1	31096	13	US-10-329-624-59	Sequence 59, Appli
74	47.8	2.1	1671	13	US-10-412-698B-1191	Sequence 1191, Ap
75	47.4	2.1	1671	16	US-10-374-780A-834	Sequence 834, App
76	47.4	2.1	536	15	US-10-029-386-19974	Sequence 19974, A
77	47.2	2.1	536	15	US-10-029-386-19974	Sequence 6243, Ap
78	47.2	2.1	599	15	US-10-029-386-6243	Sequence 6243, Ap
79	47	2.1	143899	10	US-09-972-546-15	Sequence 15, Appli
80	46.8	2.0	646	13	US-10-424-599-39641	Sequence 39641, A
81	46.6	2.0	1482	9	US-09-070-327A-842	Sequence 842, App
82	46.4	2.0	327	9	US-09-864-761-28059	Sequence 28059, A
83	46.4	2.0	439	9	US-09-960-352-1009	Sequence 1009, Ap
84	46.4	2.0	193853	13	US-10-087-192-1663	Sequence 1663, Ap
85	46.2	2.0	1282	14	US-10-002-344A-89	Sequence 89, Appli
86	46.2	2.0	2392	15	US-10-084-817-312	Sequence 312, App
87	46.2	2.0	5631	13	US-10-282-122A-37677	Sequence 37677, A

C	88	46	2.0	522	14	US-10-101-487-71
	89	46	2.0	530	14	US-10-101-487-73
	90	46	2.0	554	14	US-10-101-487-69
	91	46	2.0	554	14	US-10-101-487-106
	92	45-8	2.0	2916	13	US-10-281-122A-17169
	93	45-8	2.0	5394	15	US-10-311-455-981
	94	45-4	2.0	1848	13	US-10-021-632-98794
	95	45-4	2.0	1848	13	US-10-021-632-98795
	96	45-4	2.0	1848	16	US-10-021-632-98794
	97	45-4	2.0	1848	16	US-10-021-632-98795
	98	45-2	2.0	423	9	US-09-864-761-18355
	99	45-2	2.0	68233	16	US-10-031-650-31
	100	45-2	2.0	118502	13	US-10-081-192-397
	101	44-8	2.0	4316	9	US-09-880-107-3713
	102	44-8	2.0	4316	9	US-10-251-021-81
	C	103	44-8	2.0	14006	15
104		44-8	2.0	367378	15	US-10-312-841-2
105		44-6	1.9	1680	10	US-09-290-523B-73
106		44-6	1.9	1680	13	US-09-748-875-73
107		44-6	1.9	20142	13	US-10-282-122A-7627
108		44-4	1.9	286	15	US-10-029-386-18943
109		44-4	1.9	569	13	US-10-424-593-102311
110		44-4	1.9	577	15	US-10-029-386-5187
111		44-4	1.9	1047	13	US-10-335-977-821
112		44-4	1.9	1050	13	US-10-335-977-821
113		44-2	1.9	451	9	US-09-960-352-10262
114		44-2	1.9	462	9	US-09-864-761-3829
115		44-2	1.9	804	13	US-10-021-632-29676
116		44-2	1.9	804	16	US-10-021-632-29676
117		44-2	1.9	6568	15	US-10-311-455-1670
C		118	44-2	1.9	106664	15
	119	44	1.9	554	13	US-10-027-632-64906
	120	44	1.9	554	16	US-10-027-632-64906
	121	44	1.9	1410	9	US-09-815-242-7436
	122	44	1.9	1410	13	US-10-335-977-823
	123	44	1.9	1410	13	US-10-335-977-823
	124	44	1.9	2522	9	US-09-842-552-98
	125	44	1.9	1691139	15	US-10-087-513-1
	126	44	1.9	1691139	16	US-10-419-724-1
	127	44	1.9	367378	15	US-10-312-841-1
	128	43-8	1.9	483	9	US-09-920-300A-439
	129	43-8	1.9	483	14	US-10-031-528-439
	130	43-8	1.9	483	15	US-10-099-926-439
	131	43-8	1.9	554	9	US-09-796-632-4144
	132	43-8	1.9	554	15	US-10-040-862-4144
	133	43-8	1.9	554	16	US-10-057-475B-4144
C	134	43-8	1.9	554	16	US-10-154-884B-4144
	135	43-8	1.9	765	14	US-10-141-627-3
	136	43-8	1.9	855	14	US-10-141-627-5
	137	43-8	1.9	2520	17	US-10-389-566-205
	138	43-8	1.9	3061	9	US-09-833-790-414
	139	43-8	1.9	3061	9	US-09-880-107-3751
	140	43-8	1.9	3061	13	US-10-386-160-365
	141	43-8	1.9	3071	9	US-09-860-107-1651
	142	43-8	1.9	3253	15	US-09-811-353-20312
	143	43-8	1.9	8149	15	US-10-311-455-1123
	144	43-8	1.9	126413	13	US-10-087-192-1831
	145	43-6	1.9	5195	15	US-10-311-455-893
	146	43-6	1.9	367378	15	US-10-312-841-1
	147	43-4	1.9	767	13	US-09-424-599-7401
	148	43-4	1.9	2014	9	US-09-842-552-22
	C	149	43-4	1.9	2065	15

ALIGNMENTS

RESULT 1
US-09-765-272-65
; Sequence 65, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:

```

/ APPLICANT: Choi et. al.
/ TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
/ NUMBER OF SEQUENCES: 452
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/765,272
/ FILING DATE: 22-Jan-2001
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/961,083
/ FILING DATE: <Unknown>
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brookes, A. Anders
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: PB340P2
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/
/ INFORMATION FOR SEQ ID NO: 65:
/ SEQUENCE CHARACTERISTICS:
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 65:
/
/ US-09-765-272-65
/
/ Query Match          99.8%; Score 2285; DB 9; Length 2290;
/ Best Local Similarity 100.0%; Pred. No. 0;
/ Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY      1  TTGTTCTATGAACTTGGTTCGTCACCAAGCTGGTCAGGTTTGAAGAGTCTTAATCGAGT 60
/ Db      1  TTGTTCTATGAACTTGGTTCGTCACCAAGCTGGTCAGGTTTGAAGAGTCTTAATCGAGT 60
/
/ QY      61  TTCTTATATAGATGGTGATCAGGCTCGTCAAAAGGCAGAGAAACTTCACACCAAGTGAAGT 120
/ Db      61  TTCTTATATAGATGGTGATCAGGCTCGTCAAAAGGCAGAGAAACTTCACACCAAGTGAAGT 120
/
/ QY     121  CAGTAAAGAGGGAGGGGATCAAGCCGAAACAAATGNTGNATCAAGATTACGGATCAAGGTTA 180
/ Db     121  CAGTAAAGAGGGAGGGGATCAAGCCGAAACAAATGNTGNATCAAGATTACGGATCAAGGTTA 180
/
/ QY     181  TGTGACCTCTCATGGAGAGCCATTATCATCTACTATATAATGCAAGGTTCCCTTATCATGCCAT 240
/ Db     181  TGTGACCTCTCATGGAGAGCCATTATCATCTACTATATAATGCAAGGTTCCCTTATCATGCCAT 240
/
/ QY     241  CATTCAGTGAAGAGCTCCTCATGAAGATCCGAATTATCAGTTCAAGAGATTTCAGACATTGT 300
/ Db     241  CATTCAGTGAAGAGCTCCTCATGAAGATCCGAATTATCAGTTCAAGAGATTTCAGACATTGT 300
/
/ QY     301  CAATGAATCAAGGGTGGTTATGTCATTAAAGGTTAAACGGTAAATATCTATGNTACCTTAA 360
/ Db     301  CAATGAATCAAGGGTGGTTATGTCATTAAAGGTTAAACGGTAAATATCTATGNTACCTTAA 360
/
/ QY     361  GGATGCAGCTCATGCGGATAAATTTCGGACAAAGAAGAGATTAAACGTCAGAGCAGGA 420
/ Db     361  GGATGCAGCTCATGCGGATAAATTTCGGACAAAGAAGAGATTAAACGTCAGAGCAGGA 420
/
/ QY     421  ACGCAGTCAATAATCATAACTCAAGAGCAGATAATGCTGTGTCGACCCAGACCCCAAGG 480
/ Db     421  ACCCAGTCAATAATCATAACTCAAGAGCAGATAATGCTGTGTCGACCCAGACCCCAAGG 480

```

481 ACCTTATACACGAGTATGGGTATATCTCAATGCATCTGATATCATGAGACACGGG 540
481 ACCTTATACACGAGTATGGGTATATCTCAATGCATCTGATATCATGAGACACGGG 540
541 TGATGCTTATATCGTTCTCCACGGGACCCATTACCAATTACATTCCTTAAGAATGAGTTATC 600
541 TGATGCTTATATCGTTCTCCACGGGACCCATTACCAATTACATTCCTTAAGAATGAGTTATC 600
601 AGCTAGCGAGTTAGCTGCTGACAGAGCCCTTTGGAAATGGGAAGCAGGGATCTCGTCCCTTC 660
601 AGCTAGCGAGTTAGCTGCTGACAGAGCCCTTTGGAAATGGGAAGCAGGGATCTCGTCCCTTC 660
661 TTCAAGTCTTAGTTAATAAGTCAAAATCCAGTCAACCAAGATGTCAGAGAACCCAAATCT 720
661 TTCAAGTCTTAGTTAATAAGTCAAAATCCAGTCAACCAAGATGTCAGAGAACCCAAATCT 720
721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTACGTGA 780
721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTACGTGA 780
781 ATTGTATGCTAAACCCCTTATCAGAACGCCCTGTGGAAATCTGATGGCCTTATTTGACCC 840
781 ATTGTATGCTAAACCCCTTATCAGAACGCCCTGTGGAAATCTGATGGCCTTATTTGACCC 840
841 AGCGCAAAATCACAAGTCGAACCCGACAGCTGTAGCTGTCCCTCATGGTAACCATTTACCA 900
841 AGCGCAAAATCACAAGTCGAACCCGACAGCTGTAGCTGTCCCTCATGGTAACCATTTACCA 900
901 CTTTATCCCTTATGAACAAATGTCTGAATGGGAAACCGAATTCCTGTAATTTATCCCT 960
901 CTTTATCCCTTATGAACAAATGTCTGAATGGGAAACCGAATTCCTGTAATTTATCCCT 960
961 TCGTTATCGTTCAACCATTTGGTACAGCTGTGAGCTGTCCCTCATGTAACCAATTC 1020
961 TCGTTATCGTTCAACCATTTGGTACAGCTGTGAGCTGTCCCTCATGTAACCAATTC 1020
1021 GACTCCGGAACTTAGTCCAAAGTCCGAACTGACCAACCAATTCCTCAACCAAGTCCAAAGCAA 1080
1021 GACTCCGGAACTTAGTCCAAAGTCCGAACTGACCAACCAATTCCTCAACCAAGTCCAAAGCAA 1080
1081 TCCAAATGTATGAAATTTGGTCAAGAAGCTGTGTAAGGATAGGCGATGTTATGCTTT 1140
1081 TCCAAATGTATGAAATTTGGTCAAGAAGCTGTGTAAGGATAGGCGATGTTATGCTTT 1140
1141 TGAGGAGATGAGTTCTCGTTATATCCAGCCAGGATCTTTCAGCAGAACACGAGC 1200
1141 TGAGGAGATGAGTTCTCGTTATATCCAGCCAGGATCTTTCAGCAGAACACGAGC 1200
1201 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260
1201 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260
1261 AACTGACCTCCCATCTAGTATGATCGAATTTTACAATAAGGCTTATGACTTACTAGCAAG 1320
1261 AACTGACCTCCCATCTAGTATGATCGAATTTTACAATAAGGCTTATGACTTACTAGCAAG 1320
1321 AATTCAACCAAGATTTACTTGAATAAAGCTGACCAAGTTGATTTTGGGCTTTGGATAA 1380
1321 AATTCAACCAAGATTTACTTGAATAAAGCTGACCAAGTTGATTTTGGGCTTTGGATAA 1380
1381 CCGTTTGGAACTCAAGATGTCNCAATGTAATAAAGTCAAGTTAGTGAGATATCT 1440
1381 CCGTTTGGAACTCAAGATGTCNCAATGTAATAAAGTCAAGTTAGTGAGATATCT 1440
1441 TGCCTTTTAGCTCCGATTCGTCATCCAGACGTTTGGAAAAACCAATTCGCAAAATTC 1500
1441 TGCCTTTTAGCTCCGATTCGTCATCCAGACGTTTGGAAAAACCAATTCGCAAAATTC 1500
1501 CTACACTGATGATGATTCAGTAGGCAATTTGGGAGGCAAGTACACACAGAACGCG 1560
1501 CTACACTGATGATGATTCAGTAGGCAATTTGGGAGGCAAGTACACACAGAACGCG 1560

1561 TTATATCTTTGATCTCGTATATTAACCAAGTATGAGGGGATGCCCTATGTAATCCACA 1620
1561 TTATATCTTTGATCTCGTATATTAACCAAGTATGAGGGGATGCCCTATGTAATCCACA 1620
1621 TATGACCCATAGCCACTGGATTAAAAAGATATGTTTGTCTGAAGCTGAGAGACGCGCAGC 1680
1621 TATGACCCATAGCCACTGGATTAAAAAGATATGTTTGTCTGAAGCTGAGAGACGCGCAGC 1680
1681 CAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGATTTCAG 1740
1681 CAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGATTTCAG 1740
1741 AAATACTGAGGCAAAAGGAGCAGAGCTATCTACACCCGCTGAAAGCAGCTAAGAGGT 1800
1741 AAATACTGAGGCAAAAGGAGCAGAGCTATCTACACCCGCTGAAAGCAGCTAAGAGGT 1800
1801 GGCATTTGATCTGCTTATGCTTCAATCTCAATATCTGTAAGTCAAAAAACGTTAGTTT 1860
1801 GGCATTTGATCTGCTTATGCTTCAATCTCAATATCTGTAAGTCAAAAAACGTTAGTTT 1860
1861 AATCATACCTCATTTATGACCATTAACCAATCAATTTGAGTGGTCTGACGAGGCCCT 1920
1861 AATCATACCTCATTTATGACCATTAACCAATCAATTTGAGTGGTCTGACGAGGCCCT 1920
1921 TTATGAGGCACTTAAGGGGTATATCTCTGAGGATCTTTTGGGACCTGTCAAAGTACTATGT 1980
1921 TTATGAGGCACTTAAGGGGTATATCTCTGAGGATCTTTTGGGACCTGTCAAAGTACTATGT 1980
1981 CGAATATCCAAACGAACTGCGCATTCAGATATGTTTGGTAACTGAGCAGCAGCATGT 2040
1981 CGAATATCCAAACGAACTGCGCATTCAGATATGTTTGGTAACTGAGCAGCAGCATGT 2040
2041 TCAGAGAAACAAATATGCTCAAGCTGATACCAATCAACCGAAAAACCAAGCAGAGAA 2100
2041 TCAGAGAAACAAATATGCTCAAGCTGATACCAATCAACCGAAAAACCAAGCAGAGAA 2100
2101 ACCTCAGACAGAAAAACCTTGAGAGAAAAACCCCTCGAGAGAGAAAAACCGCAAGCAGAA 2160
2101 ACCTCAGACAGAAAAACCTTGAGAGAAAAACCCCTCGAGAGAGAAAAACCGCAAGCAGAA 2160
2161 ACAGAGTCTCCAAACCAACAGAGAACAGAGAAATCACAGAGGAATCAGAGAAC 2220
2161 ACAGAGTCTCCAAACCAACAGAGAACAGAGAAATCACAGAGGAATCAGAGAAC 2220
2221 TCAGGTCGAGACTGAAAAAGTTGAAGAAAACTGAGAGAGGCTGAAGATTACTTCGAAA 2280
2221 TCAGGTCGAGACTGAAAAAGTTGAAGAAAACTGAGAGAGGCTGAAGATTACTTCGAAA 2280
2281 AATCCAGGAT 2290
2281 AATCCAGGAT 2290

RESULT 2

US-09-769-787-206
; Sequence 206, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129W
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 206
; LENGTH: 2481

; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-206

Query Match 98.0%; Score 22.4.8; DB 10; Length 2481;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2268; Conservative 0; Mismatches 22; Indels 3; Gaps 1;

QY	1	TTGTTCCATGAACTTGCTGCTCACCAGCTGGTTCAGGTTTAAAGAAAGAGTCTAATCGAGT	60	DB	1017	CCITCGTTATCGTTGTTCAACCCATTTGGTACCAGATTCAAGACCAAGAAACAACCAAGTCCACA	1076
DB	57	TTGTTCCATGAGCTTGCTGCTCACCAGCTGGTTCAGGTTTAAAGAAAGAGTCTAATCGAGT	116	QY	1018	ATCGACTCCGGAACTTAGTCCAAAGTCGCAACCTGCAACCAATCCTCAACCAAGTCCAAAG	1077
QY	61	TTCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCGAGAAACTTGACACCAAGATGAAGT	120	DB	1077	ATCGACTCCGGAACTTAGTCCAAAGTCGCAACCTGCAACCAATCCTCAACCAAGTCCAAAG	1136
DB	117	TGCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCGAGAAACTTGACACCAAGATGAAGT	176	QY	1078	CAATCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTGCAAAAAGTAGGCGATGTTATGT	1137
QY	121	CAGTAAGAGGGGGGATCAACGCCGAAACAAATNGTNATCAAGNTTACGATCAAGTTTA	180	DB	1137	CAATCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTGCAAAAAGTAGGCGATGTTATGT	1196
DB	177	CAGTAAGAGGGGGGATCAACGCCGAAACAAATNGTNATCAAGNTTACGATCAAGTTTA	236	QY	1138	CTTTGAGGAAATGAGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAAGAACAGC	1197
QY	181	TGTGACCTCTCATGGAGACCAATATCATTAATAATGGCAAGGTTCTTTATGATGCCAT	240	DB	1197	CTTTGAGGAAATGAGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAAGAACAGC	1256
DB	237	TGTGACCTCTCATGGAGACCAATATCATTAATAATGGCAAGGTTCTTTATGATGCCAT	296	QY	1198	AGCAGGCAATTTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGAGCTAA	1257
QY	241	CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATGT	300	DB	1257	AGCAGGCAATTTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGAGCTAA	1316
DB	297	CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATGT	356	QY	1258	GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTTATGACTTACTAGC	1317
QY	301	CAATGAATCAAGGGTGGTTATGTCATTAAGTTAAAGGTAAACGGTAAATACATGTTACCTTAA	360	DB	1317	GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTTATGACTTACTAGC	1376
DB	357	CAATGAATCAAGGGTGGTTATGTCATTAAGTTAAAGGTAAACGGTAAATACATGTTACCTTAA	416	QY	1318	AGAAATTCACCAAGATTTACTTGATTAATAAGGTGCAAGATTTGATTTGAGGCTTTGGA	1377
QY	361	GGATGACGCTCATCGGATTAATATTCGGCAAAAAGAGATTAACCGTCAGAGCAGGA	420	DB	1377	AGAAATTCACCAAGATTTACTTGATTAATAAGGTGCAAGATTTGATTTGAGGCTTTGGA	1436
DB	417	GGATGACGCTCATCGGATTAATATTCGGCAAAAAGAGATTAACCGTCAGAGCAGGA	476	QY	1378	TAACCTGTTTGGAAAGCTCAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGGANGATAT	1437
QY	421	ACGAGTCATAATCAATCAACCAAGAGC...AGATAATGCTGTTGCTGCGAGCAGAGCCCA	477	DB	1437	TAACCTGTTTGGAAAGCTCAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGGANGATAT	1496
DB	477	ACGAGTCATAATCAATCAACCAAGAGCTACGATCATGTCAGTAGCTGCGAGCAGAGCCCA	536	QY	1438	TCCTTGCTCTTTAGCTCCGATTCCTCATCAGAAACGTTTATAGGAAACCAATGCGCAAT	1497
QY	478	AGGACGTTATACACGAGTATGGGTATTCCTTCAATGCACTGATATCATTGAGGACAC	537	DB	1497	TCCTTGCTCTTTAGCTCCGATTCCTCATCAGAAACGTTTATAGGAAACCAATGCGCAAT	1556
DB	537	AGGACGTTATACACGAGTATGGGTATTCCTTCAATGCACTGATATCATTGAGGACAC	596	QY	1498	TACCTACATGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACAACAGAGA	1557
QY	538	GGGTGATGCTTATATCGTTCTCAGCGCAACCATTAACATTTACATTCCTTAAGATGAGTT	597	DB	1557	TACCTACATGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACAACAGAGA	1616
DB	597	GGGTGATGCTTATATCGTTCTCAGCGCAACCATTAACATTTACATTCCTTAAGATGAGTT	656	QY	1558	CGGTTATATCTTTGATCTCGTGATTAACCAAGTGAAGGGGATGCTTATGTAATCTCC	1617
QY	598	ATCAGCTAGCGAGTTAGTCTGCTCAGAAACCTTATGGAATGGGAAGCAGGATCTCGTCC	657	DB	1617	CGGTTATATCTTTGATCTCGTGATTAACCAAGTGAAGGGGATGCTTATGTAATCTCC	1676
DB	657	ATCAGCTAGCGAGTTAGTCTGCTCAGAAACCTTATGGAATGGGAAGCAGGATCTCGTCC	716	QY	1618	ACATATGACCCATAGCCACTGGATTAAAAAGATAGTTTCTGTAAGCTCAGAGAGCGGC	1677
QY	658	TTCTTCAAGTTCTAGTTATATCAATCAACCTCAGCTCAACCAAGATTTGTCAGAGAACCAAA	717	DB	1677	ACATATGACCCATAGCCACTGGATTAAAAAGATAGTTTCTGTAAGCTCAGAGAGCGGC	1736
DB	717	TTCTTCAAGTTCTAGTTATATCAATCAACCTCAGCTCAACCAAGATTTGTCAGAGAACCAAA	776	QY	1678	AGCCAGGCTTATGCTTAAGAGAAAGTTTGAACCTCTTCGACAGACCATCAGGATTC	1737
QY	718	TCTGACTGTCACTCCAACTTATATCAAAATCAAGAGGAAACATTTCAAGCCCTTTTACG	777	DB	1737	AGCCAGGCTTATGCTTAAGAGAAAGTTTGAACCTCTTCGACAGACCATCAGGATTC	1796
DB	777	TCTGACTGTCACTCCAACTTATATCAAAATCAAGAGGAAACATTTCAAGCCCTTTTACG	836	QY	1738	AGGAAATACCTGAGGCAAAAGGAGCAGAAAGCTATCTACAACCGGCTGAAAGCAGCTAAGAA	1797
QY	778	TGAATTTGATGCTAAACCTTTATCAGAAAGCCATGTGGAATCTGATGGCCCTTATTTGGA	837	DB	1797	AGGAAATACCTGAGGCAAAAGGAGCAGAAAGCTATCTACAACCGGCTGAAAGCAGCTAAGAA	1856
DB	837	TGAATTTGATGCTAAACCTTTATCAGAAAGCCATGTGGAATCTGATGGCCCTTATTTTGA	896	QY	1798	GGTGCCACTTGTATGCTTACAAATCTTCAATATACCTGTAGAGTCAAAAACCGTAG	1857
QY	838	CCCAGCGAAATCAAGTTCGAAACCGCAAGAGTGTAGCTGCTCCTCATGTGTAAACCATTA	897	DB	1857	GGTGCCACTTGTATGCTTACAAATCTTCAATATACCTGTAGAGTCAAAAACCGTAG	1916
DB	897	CCCAGCGAAATCAAGTTCGAAACCGCAAGAGTGTAGCTGCTCCTCATGTGTAAACCATTA	956	QY	1858	TTTAAATCATACCTTATGACCAATTTACATACATCAATTTTGAAGTGTTCGACGAAGG	1917
QY	898	CCACTTTATCCCTTATGAAACCAATGTCTCAATTTGGAATAAGAAATGCTGCTATTTATTC	957	DB	1917	TTTAAATCATACCTTATGACCAATTTACATACATCAATTTTGAAGTGTTCGACGAAGG	1976
DB	957	CCACTTTATCCCTTATGAAACCAATGTCTCAATTTGGAATAAGAAATGCTGCTATTTATTC	1016	QY	1918	CTTTTATGAGGCACTTAAAGGGGTATCTCTTGAAGGATCTTTTGGCGACTGTCAAGTACTA	1977
QY	958	CCTTCGGTTATCGTTCAAAACCAATTTGGGTAACAGATTCAGAACCAACCAAGTCCACA	1017	DB	1977	CTTTTATGAGGCACTTAAAGGGGTATCTCTTGAAGGATCTTTTGGCGACTGTCAAGTACTA	2036

Db 2097 TGTTCAAGAAACAAAAATGGTCAAGTCGATACCAATCAAAACGAAAAACCAAGCGAGGA 2156
QY 2098 GAAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAAAGAGAAAAACCGCAAGCGGA 2157
Db 2157 GAAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAAAGAGAAAAACCGCAAGCGGA 2216
QY 2158 GAAACAGAGTCTCCAAACAAACAGAGGAAACAGAGGAAACAGAGAAATCAACAGAGAAATCAAGAGA 2217
Db 2217 GAAACAGAGTCTCCAAACAAACAGAGGAAACAGAGGAAACAGAGAAATCAACAGAGAAATCAAGAGA 2276
QY 2218 ACCTCAGGTCGAGACTGAAAGGTTGAAGAAACCTGAGAGAGGCTGAAGATTTACTTGG 2277
Db 2277 ACCTCAGGTCGAGACTGAAAGGTTGAAGAAACCTGAGAGAGGCTGAAGATTTACTTGG 2336
QY 2278 AAAAATCCAGGAT 2290
Db 2337 AAAAATCCAGGAT 2349

RESULT 3
US-10-412-862-11
; Sequence 11, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leelie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; PRIORITY FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 11
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-862-11

Query Match 97.0%; Score 2220.2; DB 13; Length 2531;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2253; Conservative 3; Mismatches 16; Indels 18; Gaps 1;

QY 1 TTGTTCTTAAGAACTTGGTCGTCACCAAGTTGGTCAGGTTAAGAAAGAGTCTTAATCGAGT 60
Db 57 TTGTTCTTAAGAACTTGGTCGTCACCAAGTTGGTCAGGTTAAGAAAGAGTCTTAATCGAGT 116
QY 61 TTCTTATATAGATGGTGATCAGCGTGGTCAAAAGGAGAGAACTTGAACCAAGATCGAGT 120
Db 117 TCGTTATATAGATGGTGATCAGCGTGGTCAAAAGGAGAGAACTTGAACCAAGATCGAGT 176
QY 121 CAGTAAGAGGAGGAGGATCAACCGCGAACAATGTTATCAAGATTAACGATCAAGGTTA 180
Db 177 CAGTAAGAGGAGGAGGATCAACCGCGAACAATGTTATCAAGATTAACGATCAAGGTTA 236
QY 181 TGTGACCTCTCATGGAGACCATATATCAATTAATATGGCAAGGTTCCCTTATGAGCCAT 240
Db 237 TGTGACCTCTCATGGAGACCATATATCAATTAATATGGCAAGGTTCCCTTATGAGCCAT 296
QY 241 CATCAGTGAAGAGCTCCTCATGAAAGATCGAATATCAAGTGAAGGATTCAGACATGCT 300
Db 297 CATCAGTGAAGAGCTCCTCATGAAAGATCGAATATCAAGTGAAGGATTCAGACATGCT 356
QY 301 CAATGAAATCAAGGGTGGTTATGTCATTAAGGTAAACGGTAAATCTATGTTTACCTTAA 360
Db 357 CAATGAAATCAAGGGTGGTTATGTCATTAAGGTAAACGGTAAATCTATGTTTACCTTAA 416
QY 361 GGATGAGCTCATGGGATAATATTCGGAATAAAGAGAGATTAACCGTCAGAGCGGA 420

Db 417 GGATGCRGCTCATGGGATAATATTCGGACAAAGAGAGATTAACCGTCAGAGCGAGGA 476
QY 421 AGCAGTCAATATCATAACTCAAGAGCAGATATGCTGTGTGTCAGAGCAGAGCCCAAGG 480
Db 477 AGCGAGTCAATATCATAACTCAAGAGCAGATATGCTGTGTGTCAGAGCAGAGCCCAAGG 536
QY 481 AGCTTATACAAACGAGTGAATGGGTATATCTTCAATGATCTGATATCATATTGAGACACGGG 540
Db 537 AGCTTATACAAACGAGTGAATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG 596
QY 541 TGATGCTTATATCGTTCTCAACGGGACCAATTAACCATTAATCTTAAGAAATGAGTTATC 600
Db 597 TGATGCTTATATCGTTCTCAACGGGACCAATTAACCATTAATCTTAAGAAATGAGTTATC 656
QY 601 AGCTAGCGAGTTAGTCTGCGAGAGGCTTATGGATGGAGAGGAGGATCTCGTCTTTC 660
Db 657 AGCTAGCGAGTTAGTCTGCGAGAGGCTTATGGATGGAGAGGAGGATCTCGTCTTTC 716
QY 661 TTCAAGTCTTAGTTATATGCAAAATCCAGCTCAACCAAGATTTGTCAAGAGAACCAATCT 720
Db 717 TTCAAGTCTTAGTTATATGCAAAATCCAGCTCAACCAAGATTTGTCAAGAGAACCAATCT 776
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 780
Db 777 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 836
QY 781 ATTGATGCTAAACCTTATCAGAGGCAATGGAATCTGATGGCTTATTTTGACCC 840
Db 837 ATTGATGCTAAACCTTATCAGAGCGCAATGGAATCTGATGGCTTATTTTGACCC 896
QY 841 AGCGCAAAATCAAGTCCGAACCGCGAGAGTGTAGCTGCTCCCTCATGTAAACCATTAACA 900
Db 897 AGCGCAAAATCAAGTCCGAACCGCGAGAGTGTAGCTGCTCCCTCATGTAAACCATTAACA 956
QY 901 CTTTATCCCTTATGAACAAATGCTGAAATGGAAAAACGAATTCGTCGTAATTTATCCCT 960
Db 957 CTTTATCCCTTATGAACAAATGCTGAAATGGAAAAACGAATTCGTCGTAATTTATCCCT 1016
QY 961 TCGTTATCGTTCAAAACCATTTGGTGTACAGATTCAGAGCCAGAACCAAGTCCCAATC 1020
Db 1017 TCGTTATCGTTCAAAACCATTTGGTGTACAGATTCAGAGCCAGAACCAAGTCCCAATC 1076
QY 1021 GACTCGGAAACCTAGTCCAAAGTCCGCAACCTCCACCAATCTCAACAGCTCCCAAGCAA 1080
Db 1077 GACTCCAGAACCTAGTCCAAAGTCCG-----CAACCAAGCTCCCAAGCAA 1118
QY 1081 TCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTCAAAAGTAGGCGATGGTTATGCTTT 1140
Db 1119 TCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTCAAAAGTAGGCGATGGTTATGCTTT 1178
QY 1141 TGAGGAGATGGAGTTCTCGTTATATCCAGCCAGGATCTTTTCAAGCAAGACAGCAGC 1200
Db 1179 TGAGGAGATGGAGTTCTCGTTATATCCAGCCAGGATCTTTTCAAGCAAGACAGCAGC 1238
QY 1201 AGGCATTGATAGCAAACTGGCCAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260
Db 1239 AGGCATTGATAGCAAACTGGCCAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1298
QY 1261 AACTGACCTCCCATCTAGTGATCGAGAAATTTTAAATAGGCTTATGATCTAGGAG 1320
Db 1299 AACTGACCTCCCATCTAGTGATCGAGAAATTTTAAATAGGCTTATGATCTAGGAG 1358
QY 1321 AATTCCACCAAGATTTTACTTGAATAAAGTTCGAACTGATGTTGAGGCTTTCGATAA 1380
Db 1359 AATTCCACCAAGATTTTACTTGAATAAAGTTCGAACTGATGTTGAGGCTTTCGATAA 1418
QY 1381 CCGTGTGGAACGACTCAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGGAGATATTCT 1440
Db 1419 CCGTGTGGAACGACTCAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGGAGATATTCT 1478
QY 1441 TGCCTTCTTAGTCCGATTCGTCATCCAGAGCGTTTAGGAAACCAAAATGCGCAATTTAC 1500

1479	Db	TGCTCTTTAGCTCCGATTGCTCATCAGAAAGCTTTAGGAAAAACCAATCGCAATTAC	1538
1501	QY	CTACACTGATGATGAGATTCAGGTAGCCATAGTTGGCAGGCAACTACACACAGAAGACGG	1560
1539	Db	CTACACTGATGATGAGATTCAGGTAGCCATAGTTGGCAGGCAAGTACACAGCAGAAGACGG	1598
1561	QY	TTATATCTTTGATCCTCGTGATATAACCTGTGATAGGGGGATGCCTATGTACTCCACA	1620
1599	Db	TTATATCTTTGATCCTCGTGATATAACCTGTGATAGGGGGATGCCTATGTACTCCACA	1658
1621	QY	TATGACCCATAGCCACTGGATTAAAAAATAGTTTGTCTGAAGCTCGAGAGCGCGCAGC	1680
1659	Db	TATGACCCATAGCCACTGGATTAAAAAAGATAGTTTGTCTGAAGCTCGAGAGCGCGCAGC	1718
1681	QY	CCAGGCTTATGCTAAAGAGAAAAGTTTGAAGCCCTCTCTTCGACAGACCATCAGATTTCAGG	1740
1719	Db	CCAGGCTTATGCTAAAGAGAAAAGTTTGAAGCCCTCTCTTCGACAGACCATCAGATTTCAGG	1778
1741	QY	AAATACTGAGGCAAAAGAGCAGAAAGCTTCTACAACCGCGTCAAAAGCAGCTAAGAGGT	1800
1779	Db	AAATACTGAGGCAAAAGAGCAGAAAGCTTCTACAACCGMGTGAAGACGACTAAGAGGT	1838
1801	QY	GCCACTTGATCGTATGCCTTACAATCTTCAATATCTGTGAAGAGTCAAAAACGGTAGTTT	1860
1839	Db	GCCACTTGATCGTATGCCTTACAATCTTCAATATCTGTGAAGAGTCAAAAACGGTAGTTT	1898
1861	QY	AATCATACCTCATTTGACCATTTACCATACATCAAAATTTGAGTGGTTTGAAGAGGCCT	1920
1899	Db	AATCATACCTCATTTGACCATTTACCATACATCAAAATTTGAGTGGTTTGAAGAGGCCT	1958
1921	QY	TTATGAGGCACTTAAAGGGGTATCTCTTACAGGATCTTTTCGCGACATCTCAAGTACTATGT	1980
1959	Db	TTATGAGGCACTTAAAGGGGTATCTCTTACAGGATCTTTTCGCGACATCTCAAGTACTATGT	2018
1981	QY	CGAACATCCAAAACGAAACGCTCCGATTCATATAATATGTTTGGTAAAGCTAGCACCATGT	2040
2019	Db	CGAACATCCAAAACGAAACGCTCCGATTCATATAATATGTTTGGTAAAGCTAGCACCATGT	2078
2041	QY	TCAAAGAGAACAAAATGTTCAAGCTGATCCATCAAAACGAAAAACCAAGCGAGAGAA	2100
2079	Db	TCAAAGAGAACAAAATGTTCAAGCTGATCCATCAAAACGAAAAACCAAGCGAGAGAA	2138
2101	QY	ACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAAAAACCGAAAAGCGAGAA	2160
2139	Db	ACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAAAAACCGAAAAGCGAGAA	2198
2161	QY	ACCAGAGTCTCAAAAACCAACAGAGGAAACAGAAAGATCAACAGAGGAATCAGAGAAC	2220
2199	Db	ACCAGAGTCTCAAAAACCAACAGAGGAAACAGAAAGATCACCAGAGGAATCAGAGAAC	2258
2221	QY	TCAAGTTCGAGACTGAAAAAGTTCAAGAAACACTGAGAGAGGCTGAAGATTTACTTCGAAA	2280
2259	Db	TCAAGTTCGAGACTGAAAAAGTTCAAGAAACACTGAGAGAGGCTGAAGATTTACTTCGAAA	2318
2281	QY	AATCCAGGAT	2290
2319	Db	AATCCAGGAT	2328

RESULT 4

RESUL 4
US-10-412-850-11

OS-10-412-830-11
: Sequence 11. Application US/10412850

; Sequence ID, Application US/ I04;
; Publication No. US20040001836A1

; PUBLICATION NO: US200

APPLICANT: Johnson, Leslie S.

APPLICANT: JOHNSON, LESLIE E

APPLICANT: ADAMOU, JOHN E.
TITLE OF INVENTION: VACCINE COMPOSITIONS COMPRISING STREPTOCOCCUS

DATE	TITLE OF INVENTION:	VACCINE
1950	TITLE OF INVENTION:	PNEUMONIA

;	TITLE OF INVENTION:	Pneumonia
:	TITLE OF INVENTION:	Motifs

; TITLE OF INVENTION: Motifs
 : FILE REFERENCE: 469201-686

FILE REFERENCE: 469201-686
CURRENT APPLICATION NUMBER: YIC/10/412 010

; CURRENT APPLICATION NUMBER: US/10
 ; CURRENT FILING DATE: 2003-04-14

; CURRENT FILING DATE: 2003-04-14

897 AGCGCAATCAAGTCGAAACCGCCAGAGGTTAGTGTGTCCTCCCTCATGGTAAACCATTAACA 956
901 CTTTATCCCTTATGAACAAATGTCTGAATTTGAAAAACGAATTTGCTCGTATTATTCOCCT 960
957 CTTTATCCCTTATGAACAAATGTCTGAATTTGAAAAACGAATTTGCTCGTATTATTCOCCT 1016
961 TCGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACCGAACAACCAAGTCCACAATC 1020
1017 TCGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACCGAACAACCAAGTCCACAATC 1076
1021 GACTCGGGAACCTAGTCCCAAGTCCGCAACCTGCACCAATCTCAACCGCTCCAGCAA 1080
1077 GACTCCGAACCTAGTCCCAAGTCCG-----CAACCGCTCCAGCAA 1118
1081 TCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAAGTAGGCGATGTTTATGTCTT 1140
1119 TCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAAGTAGGCGATGTTTATGTCTT 1178
1141 TGAGGAGAAATGGAGTTTCTCGTTATATCCAGCAAGGATCTTTTCAGCGAACAACAGCAGC 1200
1179 TGAGGAGAAATGGAGTTTCTCGTTATATCCAGCAAGGATCTTTTCAGCGAACAACAGCAGC 1238
1201 AGGCATTCATAGCAAACTGGCCCAAGCAGCAAGATTTATCTCATAAGCTAGGAGCTAAGAA 1260
1239 AGGCATTCATAGCAAACTGGCCCAAGCAGCAAGATTTATCTCATAAGCTAGGAACTAAGAA 1298
1261 AACTGACCTCCCATCTAGTGATCGAGAAATTTTACATAAAGGCTTATGACTTACTAGCAAG 1320
1299 AACTGACCTCCCATCTAGTGATCGAGAAATTTTACATAAAGGCTTATGACTTACTAGCAAG 1358
1321 AACTCACAAGATTTACTGTATATAAAGGTCGACAAAGTTGATTTTGGGCTTTGGATAA 1380
1359 AACTCACAAGATTTACTGTATATAAAGGTCGACAAAGTTGATTTTGGGCTTTGGATAA 1418
1381 CTTGTGGAACGACTCAAGGATGTCNCAAGTGAATAAGTCAAGTATAGTGANGATATTCT 1440
1419 CTTGTGGAACGACTCAAGGATGTCNCAAGTGAATAAGTCAAGTATAGTGANGATATTCT 1478
1441 TGCCTCTTCTAGTCCGATTCGTCATCCAGAACTTTAGGAAACCAAAATGCGCAATTTAC 1500
1479 TGCCTCTTCTAGTCCGATTCGTCATCCAGAACTTTAGGAAACCAAAATGCGCAATTTAC 1538
1501 CTACATGATGATGAGATTAAGTATAGCCAAAGTTGGCAGCAAGTACACACAGCAAGACGG 1560
1539 CTACATGATGATGAGATTAAGTATAGCCAAAGTTGGCAGCAAGTACACACAGCAAGACGG 1598
1561 TTATATCTTTGATCTCGTGATATACCGATGATGAGGGGATGCTATGTAACTCCACA 1620
1599 TTATATCTTTGATCTCGTGATATACCGATGATGAGGGGATGCTATGTAACTCCACA 1658
1621 TATGACCATAGCCACTGGATTTAAAAAGATAGTTGTCTGAGCTGAGAGCGGCGC 1680
1659 TATGACCATAGCCACTGGATTTAAAAAGATAGTTGTCTGAGCTGAGAGCGGCGC 1718
1681 CCAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCTCTCGACAGACCATCAGGATTCAGG 1740
1719 CCAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCTCTCGACAGACCATCAGGATTCAGG 1778
1741 AAATCTGAGGCAAAAGGAGCAGAGCTATCTACAAACCGCGTGAAGCAGCTAAGAGGT 1800
1779 AAATCTGAGGCAAAAGGAGCAGAGCTATCTACAAACCGCGTGAAGCAGCTAAGAGGT 1838
1801 GCCACTTCATCGTATGCTTCAATCTTCAATATCTAGAGTCAAGGCTTGAAGTGGTATTT 1860
1839 GCCACTTCATCGTATGCTTCAATCTTCAATATCTAGAGTCAAGGCTTGAAGTGGTATTT 1898
1861 AATCATCTCATTTATGACCTTACATCAATCAATCAATTTGAGTGGTGGTGGAGGCGCT 1920
1899 AATCATCTCATTTATGACCTTACATCAATCAATCAATTTGAGTGGTGGTGGAGGCGCT 1958
1921 TTATGAGGCACTTAAAGGGGTATCTCTTGGAGATCTTTTGGGCGCTGTCAAGTACTATGT 1980
1959 TTATGAGGCACTTAAAGGGGTATCTCTTGGAGATCTTTTGGGCGCTGTCAAGTACTATGT 2018

QY 1381 CGAAATCCAAACGAAACGTCGCAATTCAGATAATGTTTTGTTAAACGCTAGCGACCATGT 2040
Db 2019 CGAAATCCAAACGAAACGTCGCAATTCAGATAATGTTTTGTTAAACGCTAGCGACCATGT 2078
QY 2041 TCAAGAAACAAATAATGTCMAAGCTGATACCAATCAAAACGGAACAAACCAAGCGAGAGAA 2100
Db 2079 TCAAGAAACAAATAATGTCMAAGCTGATACCAATCAAAACGGAACAAACCAAGCGAGAGAA 2138
QY 2101 ACCTCAGACAGAAACCTGAGGAGAAACCTCTCGAGAGAGAAACCGCAAGCGAGAGAA 2160
Db 2139 ACCTCAGACAGAAACCTGAGGAGAAACCTCTCGAGAGAGAAACCGCAAGCGAGAGAA 2198
QY 2161 ACCAGAGTCTCAAAACCAACAGAGAGAACCAAGAAATCAACAGAGGAATCAGAAGAAC 2220
Db 2199 ACCAGAGTCTCAAAACCAACAGAGAGAACCAAGAAATCACCAGAGGAATCAGAAGAAC 2258
QY 2221 TCAGGTCGAGACTCAAAAGTTTGAAGAAACCTGAGAGAGGCTGAAGATTTTACTTTGAAA 2280
Db 2259 TCAGGTCGAGACTCAAAAGTTTGAAGAAACCTGAGAGAGGCTGAAGATTTTACTTTGAAA 2318
QY 2281 AATCCAGGAT 2290
Db 2319 AATCCAGGAT 2328

RESULT 5
US-10-387-783-11
; Sequence 11, Application US/10387783
; Publication No. US2004005331A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 3.0
; SEQ ID NO 11
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-387-783-11

Query Match 97.0%; Score 2220.2; DB 16; Length 2531;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2253; Conservative 3; Mismatches 16; Indels 18; Gaps 1;

QY 1 TTCTTCTATGAACTTGGTCTGCACCAAGCTGTCAGGTTAAGAAAGAGTCTAATCGAGT 60
Db 57 TTCTTCTATGAGCTTGGACGTTTACCAGGCTGTCAGGATAAGAAAGAGTCTAATCGAGT 116
QY 61 TTCTTATATAGATGGTGCATCAGGCTGTCACAAAGCGCAGAAACCTTGACACAGATGAAGT 120
Db 117 TGCTTATATAGATGGTGCATCAGGCTGTCACAAAGCGCAGAAACCTTGACACAGATGAAGT 176
QY 121 CAGTAAGAGAGGAGGATCAACCGCAACAAATGTTNATCAAGATTAACGATCAAGGTTA 180
Db 177 CAGTAAGAGAGGAGGATCAACCGCAACAAATGTTNATCAAGATTAACGATCAAGGTTA 236
QY 181 TGTGACCTCTATGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Db 237 TGTGACCTCTATGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 296
QY 241 CATCAGTGAAGAGCTCTCTCATGAAGATCCGAATTAATTAATTAATTAATTAATTAATTAAT 300

Db 297 CACAGTGAAGAGCTCCTCATGAAGATCCGAATTATCAGTTGAGGATTCAGCAATTGT 356
QY 301 CAATGAATCAAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGNTACTTTAA 360
Db 357 CAATGAATCAAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGTTACTTTAA 416
QY 361 GGATGCAAGCTCATCGGATATATTCGACAAAGAGAGATTAAACGTCAAGACGAGA 420
Db 417 GGATGCGGCTCATCGGATATATTCGACAAAGAGAGATTAAACGTCAAGACGAGA 476
QY 421 ACGCAGTCATAATCATTAACCTCAAGAGCAGATAATGCTGTGTCAGCCAGAGCCCAAGG 480
Db 477 ACGCAGTCATAATCATTAACCTCAAGAGCAGATAATGCTGTGTCAGCCAGAGCCCAAGG 536
QY 481 ACGTTATACAAACGGATGATGGGTATATCTCAATGTCATCTCATATCAATTAAGATGATTC 600
Db 537 ACGTTATACAAACGGATGATGGGTATATCTCAATGTCATCTCATATCAATTAAGATGATTC 656
QY 541 TGATGCTTATATCGTTCCTCAACGGGACCATTAACCAITTAATTCCTAAGATGATTC 600
Db 597 TGATGCTTATATCGTTCCTCAACGGGACCATTAACCAITTAATTCCTAAGATGATTC 656
QY 601 AGCTAGCGAGTTAGCTGTCGAGAGCCCTATTTGGAATGGGAGCAGGGATCTCGTCCCTTC 660
Db 657 AGCTAGCGAGTTAGCTGTCGAGAGCCCTATTTGGAATGGGAGCAGGGATCTCGTCCCTTC 716
QY 661 TTCAAGTTCTAGTTATATGTCGAGAGCCCTATTTGGAATGGGAGCAGGGATCTCGTCCCTTC 720
Db 717 TTCAAGTTCTAGTTATATGTCGAGAGCCCTATTTGGAATGGGAGCAGGGATCTCGTCCCTTC 776
QY 721 GACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACCAATTTCAAGCCTTTTACGTGA 780
Db 777 GACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACCAATTTCAAGCCTTTTACGTGA 836
QY 781 ATTGTAATGCTTAAACCCCTTATCAGAACCCATGTCGAATCTGATGGCCCTTATTTTCGACC 840
Db 837 ATTGTAATGCTTAAACCCCTTATCAGAACCCATGTCGAATCTGATGGCCCTTATTTTCGACC 896
QY 841 AGCGCAATCACAACTGTCGAGAGCCGAGAGGTGATGTCCTCATGTTAAACCATTAACA 900
Db 897 AGCGCAATCACAACTGTCGAGAGCCGAGAGGTGATGTCCTCATGTTAAACCATTAACA 956
QY 901 CTTTATCCCTTATGAAACAAATGCTGAATTCGAAACCAAAATGCTGATTTATTTCCCT 960
Db 957 CTTTATCCCTTATGAAACAAATGCTGAATTCGAAACCAAAATGCTGATTTATTTCCCT 1016
QY 961 TCGTTATCGTTCAAAACCATTTGGGTACAGATTCAGACCAAGCAACCAAGTCCCAATC 1020
Db 1017 TCGTTATCGTTCAAAACCATTTGGGTACAGATTCAGACCAAGCAACCAAGTCCCAATC 1076
QY 1021 GACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAAAATCCTCAACCCAGCTCCCAAGCAA 1080
Db 1077 GACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAAAATCCTCAACCCAGCTCCCAAGCAA 1118
QY 1081 TCCAAATTGATGAGAAATGCTCAAGAGCTGTTTCGAAAGTGAAGCGATGTTATGCTT 1140
Db 1119 TCCAAATTGATGAGAAATGCTCAAGAGCTGTTTCGAAAGTGAAGCGATGTTATGCTT 1178
QY 1141 TGAGGAGATGAGGTTTCTCGTTATATCCGACCAAGATCTTTCCAGCAGAAACAGCAGC 1200
Db 1179 TGAGGAGATGAGGTTTCTCGTTATATCCGACCAAGATCTTTCCAGCAGAAACAGCAGC 1238
QY 1201 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAGTTTATCTCATTAAGCTAGGAGCTTAAGAA 1260
Db 1239 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAGTTTATCTCATTAAGCTAGGAGCTTAAGAA 1298
QY 1261 AACTGACCTCCCATCTAGTGAAGAGATTTTACAATAAGGCTTATGACTTTACTAGCAAG 1320
Db 1299 AACTGACCTCCCATCTAGTGAAGAGATTTTACAATAAGGCTTATGACTTTACTAGCAAG 1358
QY 1321 AATTCAACAGATTTACTTGATATAAAGGTGCGCAAGTTGATTTGAGGCTTTGGATAA 1380
Db 1359 AATTCAACAGATTTACTTGATATAAAGGTGCGCAAGTTGATTTGAGGCTTTGGATAA 1418

QY 1381 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATATTCT 1440
Db 1419 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATATTCT 1478
QY 1441 TGCCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAACCAAAATCGCNAATTAC 1500
Db 1479 TGCCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAACCAAAATCGCNAATTAC 1538
QY 1501 CTACACTGATGATGAGATTCAAGTAGCTGAGGCAAGTGTACACAAACAGAAAGCGG 1560
Db 1539 CTACACTGATGATGAGATTCAAGTAGCTGAGGCAAGTGTACACAAACAGAAAGCGG 1598
QY 1561 TTATATCTTTGATCTCTGATATAACCAAGTGTAGGGGGATGCTTATGTAACTCCACA 1620
Db 1599 TTATATCTTTGATCTCTGATATAACCAAGTGTAGGGGGATGCTTATGTAACTCCACA 1658
QY 1621 TATGACCCATAGGACTCGGATTAAGAAAGATGATTTGTCTGAAGCTGAGAGCGGAGC 1680
Db 1659 TATGACCCATAGGACTCGGATTAAGAAAGATGATTTGTCTGAAGCTGAGAGCGGAGC 1718
QY 1681 CCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCATCAGGATTCAGG 1740
Db 1719 CCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCATCAGGATTCAGG 1778
QY 1741 AAATACTGAGGCAAAAGAGAGCAAGCTATCTAACAACCGCTGAAAGCAGCTAAGAAAGGT 1800
Db 1779 AAATACTGAGGCAAAAGAGAGCAAGCTATCTAACAACCGMTGAAAGCAGCTAAGAAAGGT 1838
QY 1801 GGCATTGATGCTATGCTTAACTCAATCTTCAATATCTGTAGAAAGTCAAAAACGGTAGTTT 1860
Db 1839 GGCATTGATGCTATGCTTAACTCAATCTTCAATATCTGTAGAAAGTCAAAAACGGTAGTTT 1898
QY 1861 AATCATACCTCATTAAGCAATTTACCATTAACCATTAACCATTAACCATTAACCATTAACCAT 1920
Db 1899 AATCATACCTCATTAAGCAATTTACCATTAACCATTAACCATTAACCATTAACCATTAACCAT 1958
QY 1921 TTATGAGGCACTTAAGGGGTATATCTCTGAGGATCTTTTGGCGACTGTCAAGTACTATGT 1980
Db 1959 TTATGAGGCACTTAAGGGGTATATCTCTGAGGATCTTTTGGCGACTGTCAAGTACTATGT 2018
QY 1981 CGAAATCAAAACGAAACGTCGCAATTCAGATTAATGTTTGGTAAACGCTAGCGACCATGT 2040
Db 2019 CGAAATCAAAACGAAACGTCGCAATTCAGATTAATGTTTGGTAAACGCTAGCGACCATGT 2078
QY 2041 TCAAGAAACAAAAATGCTCAAGCTGATACCATTAACCGAAACCAACCAACCAACCAACCAAC 2100
Db 2079 TCAAGAAACAAAAATGCTCAAGCTGATACCATTAACCGAAACCAACCAACCAACCAACCAAC 2138
QY 2101 ACCTCAGACAGAAACCTGAGGAGAAACCCCTCGAGAGAGAAACCCGAAACCGAGAGAA 2160
Db 2139 ACCTCAGACAGAAACCTGAGGAGAAACCCCTCGAGAGAGAAACCCGAAACCGAGAGAA 2198
QY 2161 ACCAGAGTCTCAAAACCAACAGAGAAACCAAGAAATCAGAGAGAAATCAGAGAAAC 2220
Db 2199 ACCAGAGTCTCAAAACCAACAGAGAAACCAAGAAATCAGAGAGAAATCAGAGAAAC 2258
QY 2221 TCAGTTCAGAGCTGAAAGGTTGAAGAAACCACTGAGAGAGGCTGAAGATTTACTTGA 2280
Db 2259 TCAGTTCAGAGCTGAAAGGTTGAAGAAACCACTGAGAGAGGCTGAAGATTTACTTGA 2318
QY 2281 AATCCAGGAT 2290
Db 2319 AATCCAGGAT 2328

RESULT 6

US-09-884-465A-5
; Sequence 5, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josée

APPLICANT: Brodeur, Bernard
APPLICANT: Martin, Denis
APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 055190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5

LENGTH: 2639

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

JS-09-884-465A-5

Query Match 88.3%; Score 2022.2; DB 10; Length 2639;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 148; Indels 24; Gaps 1;

1 TTGTTCCCTATGAACCTTGGTCGTCACCAAGCTGTCAGGTTAAGAGAGCTCTAATCGAGT 60
170 TTGTTCCCTATGAACCTTGGTCGTCACCAAGCTGTCAGGTTAAGAGAGCTCTAATCGAGT 229

61 TTCTTATATAGATGGTGTATCAGGCTGGTCAAAAGGCGAGAAAACCTTGACACCCAGATGAAGT 120
230 TTCTTATATAGATGGTGTATCAGGCTGGTCAAAAGGCGAGAAAACCTTGACACCCAGATGAAGT 289

121 CAGTAAGAGGAGGGGGATCAACCGCGAAACAAATNGNATCAAGATTAACGATCAAGGTTA 180
290 CAGTAAGAGAGAGGGGGATCAACCGCGAAACAAATNGNATCAAGATTAACGATCAAGGTTA 349

181 TGTGACCTCTCATGAGACCATTAATCATTTACTATATATGCGAAGGTTCCCTATGATGCCAT 240
350 TGTGACCTCTCATGAGACCATTAATCATTTACTATATATGCGAAGGTTCCCTATGATGCCAT 409

241 CATCAGTGAAGAGCTCCCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTGT 300
410 CATCAGTGAAGAGCTCCCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTGT 469

301 CAATGAATCAAGGGTGGTATGTCATTAAAGTAAACCGTAAATATCATGTGTTTACCTTAA 360
470 CAATGAATCAAGGGTGGTATGTCATTAAAGTAAACCGTAAATATCATGTGTTTACCTTAA 529

361 GGATGCGAGCTCATGCGGATATATTCGGCAAAAGAGAGATTAACGTCAGAGCAGGA 420
530 AGATGCGGCGCCATGCGGACAAATATTCGGCAAAAGAGAGATTAACGTCAGAGCAGGA 589

421 ACGCAGTCATATCAATACTCAAGAGCAGATTAATGCTGTTGCTGCGAGCCAGAGCCCAAGG 480
590 ACA CAGTCATATCAATACTCAAGAGCAGATTAATGCTGTTGCTGCGAGCCAGAGCCCAAGG 649

481 ACGTTATACAAACGAGATGGGTATATCTTCAATGCATCTGATATCATTTGAGGACACGGG 540
650 ACGTTATACAAACGAGATGGGTATATCTTCAATGCATCTGATATCATTTGAGGACACGGG 709

541 TGATGCTTATATCTGTTCTTCCACGGGCAACATTAACATTAACATTCCTTAAGATGAGTTATC 600
710 TGATGCTTATATCTGTTCTTCCACGGGCAACATTAACATTAACATTCCTTAAGATGAGTTATC 769

601 AGTAGCCAGTTAGTGTGTCAGAGAGCTTATGGAATGGAGAGCAGGATCTCGTCTTTC 860
770 AGTAGCCAGTTAGTGTGTCAGAGAGCTTATGGAATGGAGAGCAGGATCTCGTCTTTC 829

661 TTCAAGTTCTAGTTTATATGCAAAATCCAGCTCAACCAAGATTTGTGAGAGAACCACAATCT 720
830 TTCAAGTTCTAGTTTATATGCAAAATCCAGTTCAACCAAGATTTGTGAGAGAACCACAATCT 889

721 GACTGTCTACTCAACTTATATCAAAAATCAAGGGGAAACATTTCAAGCTTTTACGTGA 780
890 GACTGTCTACTCAACTTATATCAAAAATCAAGGGGAAACATTTCAAGCTTTTACGTGA 949

QY 781 ATTGTATGCTAAACCCCTTATCAGAAACGCCATGTGAAATCTGATGCGCTTATTTTCGACCC 840
DB 950 ATTGTATGCTAAACCCCTTATCAGAAACGCCATGTGAAATCTGATGCGCTTATTTTCGACCC 1009

QY 841 AGCGCAATCACAAGTGCAGACCGCCAGAGGTGTAGCTGTCCCTCATGTGTAACCATTAACA 900
DB 1010 AGCGCAATCACAAGTGCAGACCGCCAGAGGTGTAGCTGTCCCTCATGTGTAACCATTAACA 1069

QY 901 CTTTATCCCTTATGCAACAAATGCTGAAATTCGAAAAACGAATTCGCTGATTTATTTCCCT 960
DB 1070 CTTTATCCCTTATGCAACAAATGCTGAAATTCGAAAAACGAATTCGCTGATTTATTTCCCT 1129

QY 961 TCGTTATTCGTTCAAAACCAATTCGGGTACAGATTCAGACCCAGAACCAACCAAGTCCACAATC 1020
DB 1130 TCGTTATTCGTTCAAAACCAATTCGGGTACAGATTCAGACCCAGAACCAACCAAGTCCACAATC 1189

QY 1021 GACTCCGGAACTTAGTCCAAAGTCGCGAAACCTGACCCAAATCCTCAACCCAGTCCCAAGAA 1080
DB 1190 GACTCCGGAACTTAGTCCAAAGTCGCGAAACCTGACCCAAATCCTCAACCCAGTCCCAAGAA 1249

QY 1081 TCCAAATTCGATGAGAAATTCGTCAAAGAGCTGTTTCGAAAAAGTAGGCGATGTTATGTCTT 1140
DB 1250 TCCAAATTCGATGAGAAATTCGTCAAAGAGCTGTTTCGAAAAAGTAGGCGATGTTATGTCTT 1309

QY 1141 TGAGGGAATTCGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTCAGCAGAAACAGCAGC 1200
DB 1310 TGAGGGAATTCGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTCAGCAGAAACAGCAGC 1369

QY 1201 AGCAATTCGATGAGAAATTCGCGCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTTAAGAA 1260
DB 1370 AGCAATTCGATGAGAAATTCGCGCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTTAAGAA 1429

QY 1261 AACTGACCTCCCATCTAGTATCGAGAAATTTTACAATAGGCTTATGACTTACTAGCAAG 1320
DB 1430 AACTGACCTCCCATCTAGTATCGAGAAATTTTACAATAGGCTTATGACTTACTAGCAAG 1489

QY 1321 AATTCAACCAAGATTTACTTGTATATAAAGGTCGACAAAGTTCATTTTGAGGCTTTTCGATAA 1380
DB 1490 AATTCAACCAAGATTTACTTGTATATAAAGGTCGACAAAGTTCATTTTGAGGCTTTTCGATAA 1549

QY 1381 CCTGTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATTTCT 1440
DB 1550 CCTGTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATTTCT 1609

QY 1441 TGGCTTCTTAGCTCCGATTCGTCATCCAGAAAGTTCGAGAAAAACCAATTCGCAAAATTTAC 1500
DB 1610 TGGCTTCTTAGCTCCGATTCGTCATCCAGAAAGTTCGAGAAAAACCAATTCGCAAAATTTAC 1669

QY 1501 CTACACTGATGATGAGATTCAGAGTAGCCAAAGTTCGCGAGCAAGTACCAACAGAGACGG 1560
DB 1670 CTACACTGATGATGAGATTCAGAGTAGCCAAAGTTCGCGAGCAAGTACCAACAGAGACGG 1729

QY 1561 TTATATCTTTGATCTCTCGTATATAACAGTGTAGAGGGGATGCTATGTAATCTCCACA 1620
DB 1730 TTATATCTTTGATCTCTCGTATATAACAGTGTAGAGGGGATGCTATGTAATCTCCACA 1789

QY 1621 TATGACCCATAGCCACTGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGCGGCGAGC 1680
DB 1790 TATGACCCATAGCCACTGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGCGGCGAGC 1849

QY 1681 CCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCATCAGGATTCAGG 1740
DB 1850 CCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCATCAGGATTCAGG 1909

QY 1741 AAATACTGAGCAAAAGAGCAGAGAGCTATCTACAACCGCGTGAAGAGCAGCTAAGAGCT 1800
DB 1910 AAATACTGAGCAAAAGAGCAGAGAGCTATCTACAACCGCGTGAAGAGCAGCTAAGAGCT 1969

QY 1801 GCCACTTTGATCGTATGCTTTTCAATCTTCAATATCTGTAGAGTCAAAAAACGGTAGTTT 1860
DB 1970 GCCACTTTGATCGTATGCTTTTCAATCTTCAATATCTGTAGAGTCAAAAAACGGTAGTTT 2029

```
QY 1861 AATCATACCTCATTATGACCAATTACCAATACCAATCAAAATTTGAGTGGTTTGACGAAGCCT 1920
DB 2030 AATCATACCTCATTATGACCAATTACCAATACCAATCAAAATTTGAGTGGTTTGACGAAGCCT 2089
QY 1921 TTATGAGGACCACTAAAGGGGTATCTCTTGAGGATCTTTTGGCGACTCTCAAGTACTATGT 1980
DB 2090 TTATGAGGACCACTAAAGGGGTATAGTCTTGAGGATCTTTTGGCGACTCTCAAGTACTATGT 2149
QY 1981 CGAATATCCAAACGAACGCTCCGCAATTCAGATAATGTTTGGTAAACGCTAGCGACCATGT 2040
DB 2150 CGAATATCCAAACGAACGCTCCGCAATTCAGATAATGTTTGGTAAACGCTAGTGAACATGT 2209
QY 2041 TCAAGAAACAAATAATGCTCAAGCTGATCCATCAACCAACGAACCAACGAGGAGAA 2100
DB 2210 TCGTAAATAAAGGCGAGACCAAGATAGTAAACCTGATGAAGAAACATGATGAAGT 2269
QY 2101 ACTCTCAGACAGAAAAACCTCAGGAAGA-----AACCCCTCG 2136
DB 2270 AAGTGAGCCAACTCACCTCGAATCTGATGAAAAGAGAATCAGCTGGTTTAAATCCTTC 2329
QY 2137 AGAAGAGAAACCGCAAGCGAGAGAAACAGAGTCTCCAAAACCAACAGAGGAACCAAGA 2196
DB 2330 AGCAGATAATCTTTATAAACCAAGCACTGATACGGAAGACAGAGGAAGAGCTGAGA 2389
QY 2197 ATCCACAGAGGAATCAGAGAACTCAGGTCGAGACTGGAAGGTTGAAAGAAACTGAG 2256
DB 2390 TACCACAGATGAGGCTGAAATTCCTCAAGTAGAGAATCTGTATTAAACGCTAAGATAGC 2449
QY 2257 AGAGGCTGAAGATTTACTTGGAAAAATCCAGGAT 2290
DB 2450 AGATCGGAGGCTTGCTAGAAAAAGTAAACAGAT 2483

RESULT 7
US-10-412-862-5
; Sequence 5, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Adamou, John E.
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; PRIORITY FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-862-5

Query Match 87.3%; Score 1999.6; DB 13; Length 2531;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 154; Indels 27; Gaps 2;

QY 1 TTGTTCTATGAACCTGGTCGTCCACCAAGCTGCTCAGGTTTAAAGAGGCTCAATCGAGT 60
DB 57 TTGTTCTATGAACCTGGTCGTCCACCAAGCTGCTCAGGTTTAAAGAGGCTCAATCGAGT 116
QY 61 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAAACTTGACACAGATGAAGT 120
DB 117 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAAACTTGACACAGATGAAGT 176
QY 121 CAGTAAAGGGAGGGGATCAACGCCGAAACAAATGNTGATCAAGATTCAGGATCAAGGTTA 180
DB 177 CAGTAAAGGGAGGGGATCAACGCCGAAACAAATCGTCAATCAAGATTCAGGATCAAGGTTA 236
```

```
QY 181 TGTGACCTCTCATGAGACCAATTATCATTTATATATGCAAGGTTTCTTTATGATGCCAT 240
DB 237 TGTGACCTCTCATGAGACCAATTATCATTTATATATGCAAGGTTTCTTTATGATGCCAT 296
QY 241 CATCATGAGAGAGCTCTCATGAAAAGATCCGAAATATCATGTTGAAGGATTCAGACATTTGT 300
DB 297 CATCATGAGAGAGCTCTCATGAAAAGATCCGAAATATCATGTTGAAGGATTCAGACATTTGT 356
QY 301 CAATGAATCAAGGGTGTATGTCATTAAGGTAAACGGTAAATCTATGTTATACCTTAA 360
DB 357 CAATGAATCAAGGGTGTATGTTATCAAGGTAGATGGAATACTATGTTTACCTTAA 416
QY 361 GGATGCAGCTCATGCGGATAATATTCGACAAAAAGAGAGATTAAACGTCAGAAAGCAGGA 420
DB 417 GGATGCAGCTCATGCGGATAATATTCGACAAAAAGAGAGATTAAACGTCAGAAAGCAGGA 476
QY 421 AGCAGTCTAATATCAATACTCAAGAGC--AGATAATGCTGTGTGCGAGCCAGAGCCCA 477
DB 477 ACAGAGTCAATAATCACGGGGGTGGTTCTAAACGATCAAGCAGTAGTTGACGCCAGAGCCCA 536
QY 478 AGGAGGTTATCAACGGGATGATGGTATATCTTCAATGCATCTGATATCATTTGAGAGCAC 537
DB 537 AGGAGGTTATCAACGGGATGATGGTATATCTTCAATGCATCTGATATCATTTGAGAGCAC 596
QY 538 GGGTGAATGCTTATATCGTTCTCACGGCGACCATTAACATTAATTCCTTAAGAAATGAGTT 597
DB 597 GGGTGAATGCTTATATCGTTCTCACGGCGACCATTAACATTAATTCCTTAAGAAATGAGTT 656
QY 598 ATCAGTAGCGAGTTAGTCTGTCAGAAAGCCTATTGGAATGGGAAGCAGGATCTCGTCC 657
DB 657 ATCAGTAGCGAGTTAGTCTGTCAGAAAGCCTATTGGAATGGGAAGCAGGATCTCGTCC 716
QY 658 TTCTTCAAGTTCTAGTTATATGAAATCCAGCTCAACCAAGATTTGTCAGAGAACACAA 717
DB 717 TTCTTCAAGTTCTAGTTATATGAAATCCAGCTCAACCAAGATTTGTCAGAGAACACAA 776
QY 718 TCTGACTGTCACTCCAACTTATCAAAATCAAGGGGAAAAACATTTCAAGCTTTTACG 777
DB 777 TCTGACTGTCACTCCAACTTATCAAAATCAAGGGGAAAAACATTTCAAGCTTTTACG 836
QY 778 TGAATTTGATGCTAAACCCCTTATCAGAACGCCATGTGAAATCTGATGGCTTTATTTTGA 837
DB 837 TGAATTTGATGCTAAACCCCTTATCAGAACGCCATGTGAAATCTGATGGCTTTATTTTGA 896
QY 838 CCCAGCGCAATCAAGTCCAAACCGCCAGAGGTGTAGTCTCCCTCATGTTAAACCATTA 897
DB 897 CCCAGCGCAATCAAGTCCAAACCGCCAGAGGTGTAGTCTCCCTCATGTTAAACCATTA 956
QY 898 CCATTTTATCCCTTATGAAACAAATGCTGTAATTTGAAAAACGAAATGCTGTTATTATTC 957
DB 957 CCATTTTATCCCTTATGAAACAAATGCTGTAATTTGAAAAACGAAATGCTGTTATTATTC 1016
QY 958 CTTTCTGTTATGTTTCAAAACCATTTGGGTACAGATTCAGAACCCAGAACCAAGTCCACA 1017
DB 1017 CTTTCTGTTATGTTTCAAAACCATTTGGGTACAGATTCAGAACCCAGAACCAAGTCCACA 1076
QY 1018 ATCGACTCCGGAACCTAGTCCAAAGTCCGCAACCTGCAACCAATTCCTCAACAGCTCCAG 1077
DB 1077 ATCGACTCCGGAACCTAGTCCAAAGTCCGCAACCTGCAACCAATTCCTCAACAGCTCCAG 1136
QY 1078 CAATCCAAATTTGATGAGAAATTTGGTCAAGAGAGCTGTTGAAAAGTAGGCGATGGTTATGT 1137
DB 1137 CAATCCAAATTTGATGAGAAATTTGGTCAAGAGAGCTGTTGAAAAGTAGGCGATGGTTATGT 1196
QY 1138 CTTTGGAGGAATGAGGTTTCTCGTTATATCCCGACCAAGGATCTTTTCAAGAGAACACAGC 1197
DB 1197 CTTTGGAGGAATGAGGTTTCTCGTTATATCCCGACCAAGGATCTTTTCAAGAGAACACAGC 1256
QY 1198 AGCAGGCAATTTGATAGCAAACTGGCCCAAGAGGAAAGTTTATCTCATAGCTAGGAGCTAA 1257
DB 1257 AGCAGGCAATTTGATAGCAAACTGGCCCAAGAGGAAAGTTTATCTCATAGCTAGGAGCTAA 1316
```


658 TTCTTCAAGTTCTAGTTTATATGCAATTCAGCTCAACCAAGATTGTTCAGAGAACCAAA 717
717 TTCTTCAAGTTCTAGTTTATATGCAATTCAGCTCAACCAAGATTGTTCAGAGAACCAAA 776
718 TCTGACTGTCACTCCCACTTATCATCAAAATCAAGGGGAAACATTTCAGCCCTTTTACG 777
777 TCTGACTGTCACTCCCACTTATCATCAAAATCAAGGGGAAACATTTCAGCCCTTTTACG 836
778 TGAATTGTATGCTAAACCCCTTATCAGAACCGCATGTGGAATCTGATGGCCCTTATTTTGA 837
837 TGAATTGTATGCTAAACCCCTTATCAGAACCGCATGTGGAATCTGATGGCCCTTATTTTGA 896
838 CCAGCGCAAAATCACAAGTGCAGACCGCGAGAGTGTAGCTGCTCCTCATGTGTACCAATTA 897
897 CCCAGCGCAAAATCACAAGTGCAGACCGCGAGAGTGTAGCTGCTCCTCATGTGTACCAATTA 956
898 CCACCTTTATCCCTTATGAAAGATGCTGTAATTCGAAAGAAAGAAATTCGCTGATTTATTC 957
957 CCACCTTTATCCCTTATGAAAGATGCTGTAATTCGAAAGAAAGAAATTCGCTGATTTATTC 1016
958 CTTTCGTTATCGTTTCAAAACCATTCGGGTACCAAGATTCAGAGCAGAAACCAACCAAGTCCACA 1017
1017 CTTTCGTTATCGTTTCAAAACCATTCGGGTACCAAGATTCAGAGCAGAAACCAACCAAGTCCACA 1076
1018 ATCGACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAAAATCCTCAACAGCTCCAAAG 1077
1077 ATCGACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAAAATCCTCAACAGCTCCAAAG 1136
1078 CAATCCAAATTCAGAGAAATGTTTCAAGAGCTGTTTCGAAAGTAGGCGATGTTATGT 1137
1137 CAATCCAAATTCAGAGAAATGTTTCAAGAGCTGTTTCGAAAGTAGGCGATGTTATGT 1196
1138 CTTTCGAGGAATGAGCTTCTCGTTATATCCAGCAGCAAGATCTTTCAGCAGAAACAGC 1197
1197 CTTTCGAGGAATGAGCTTCTCGTTATATCCAGCAGCAAGATCTTTCAGCAGAAACAGC 1256
1198 AGCAGCAATTCAGAGCAAACTGGCCAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTAA 1257
1257 AGCAGCAATTCAGAGCAAACTGGCCAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTAA 1316
1258 GAAACCTGACTCCCATCTAGTGTATGAGATTTTACAAATGAGCTTATGACTTACTAGC 1317
1317 GAAACCTGACTCCCATCTAGTGTATGAGATTTTACAAATGAGCTTATGACTTACTAGC 1376
1318 AAGAAATTCACCAAGATTACTTGTATATAAAGTTCGACAGTTCATTTTGGAGCTTTTGA 1377
1377 AAGAAATTCACCAAGATTACTTGTATATAAAGTTCGACAGTTCATTTTGGAGCTTTTGA 1436
1378 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGAGATAT 1437
1437 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGAGATAT 1496
1438 TCTTGCCTTCTTAGCTCCGATTCGATCCGAAAGCTTTAGGAAACCAATTCGCAAT 1497
1497 TCTTGCCTTCTTAGCTCCGATTCGATCCGAAAGCTTTAGGAAACCAATTCGCAAT 1556
1498 TACCTACACTGATGATGAGATTCAAGTAGCCCAAGTTGGCGAGCAAGTACACAAACAGAGA 1557
1557 TACCTACACTGATGATGAGATTCAAGTAGCCCAAGTTGGCGAGCAAGTACACAAACAGAGA 1616
1558 CGGTATATCTTTGATCCTCGTGATATAACCAAGTGTAGGGGGATGCTTATGTAATCC 1617
1617 CGGTATATCTTTGATCCTCGTGATATAACCAAGTGTAGGGGGATGCTTATGTAATCC 1676
1618 ACATATGACCCATAGCCACTGGATTAAAGATAGTTTGTGTAAGCTGAGAGCGGC 1677
1677 ACATATGACCCATAGCCACTGGATTAAAGATAGTTTGTGTAAGCTGAGAGCGGC 1736
1678 AGCCAGGCTTATGCTAAAGAGAAAGTTTGGACCCCTCCTTCGACAGCAATTCAGGATTC 1737
1737 AGCCAGGCTTATGCTAAAGAGAAAGTTTGGACCCCTCCTTCGACAGCAATTCAGGATTC 1796
1738 AGGAAATACTGAGGCAAAAGAGCAGAGCTTATCTACAAACCGCTGAAAGCAGCTAAGAA 1797

Db 1797 AGGAAATACTGAGGCAAAAGAGCAGAGCTATCTACAAACCGCTGAAAGCAGCTAAGAA 1856
QY 1798 GGTGCACTTGTATGCTTATGCTTACAAATCTTCAATATCTGTAGAAGTCAAAAACGGTAG 1857
Db 1857 GGTGCACTTGTATGCTTATGCTTACAAATCTTCAATATCTGTAGAAGTCAAAAACGGTAG 1916
QY 1858 TTTAATCATACCTCATATATGACCATTTACCAATCAATCAATTTAGTGGTTTACGAGG 1917
Db 1917 TTTAATCATACCTCATATATGACCATTTACCAATCAATCAATTTAGTGGTTTACGAGG 1976
QY 1918 CTTTATGAGGCACTTACGGGTATCTCTTGGAGTCTTTTGGCAGCTGTCAAGTACTA 1977
Db 1977 CTTTATGAGGCACTTACGGGTATCTCTTGGAGTCTTTTGGCAGCTGTCAAGTACTA 2036
QY 1978 TGTCCAAATCATCAACAGCAAGCTCCGATTCAGATTCAGATATGTTTGGTAAACGCTAGCGACCA 2037
Db 2037 TGTCCAAATCATCAACAGCAAGCTCCGATTCAGATTCAGATATGTTTGGTAAACGCTAGCGACCA 2096
QY 2038 TGTTCAAAGAAACAAAAATGTTTCAAGCTGATACCAATCAACGGAACCAAGCGAGGA 2097
Db 2097 TGTTCGTAATAAAGGTAGACCAAGACAGTAAACCTGTATGAAGATAAGGAACATGATGA 2156
QY 2098 GAAACCTCAGACAGAAAAAACCCTGAGGAAGA-----AACCC 2133
Db 2157 AGTAAAGTGAAGCAACTCAGCTGATCTGATGAAAAAGAGATCAGCTGTTTAAATCC 2216
QY 2134 TCGAGAAGAGAAACCGCAAGCGAGAAACAGTCTCCAAACCAACAGAGAGACCGA 2193
Db 2217 TTCAGCAGATATCTTTTAAACCAAGCACTGATACGGAAGAGACAGAGGAAGAAGCTGA 2276
QY 2194 AGAATCACCAGAGGAATCAGAAAGAACCTCAGGTTCGAGACTGAAAAAGTTGAAAGAAAACT 2253
Db 2277 AGATACCAAGATGAGGTGAAATTCCTCAGTAGAGATTCGTTTATTAACGTAAGAT 2336
QY 2254 GAGAGAGCTGAAGATTTACTTTGGAAAAATCCAGGAT 2290
Db 2337 AGCAGATGGGAGGCTTGTCTAGAAAAAGTAACAGAT 2373

RESULT 9
US-10-387-783-5
; Sequence 5, Application US/10387783
; Publication No. US20040005331A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 5
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-387-783-5

Query Match 87.3%; Score 1999.6; DB 16; Length 2531;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 154; Indels 27; Gaps 2;
QY 1 TTGTTCTTATGAATTTGGTGTGTCACCAAGCTGGTCAAGGTTAAGAAAGTCTTAATCAGT 60
Db 57 TTGTTCTTATGAATTTGGTGTGTCACCAAGCTGGTCAAGGTTAAGAAAGTCTTAATCAGT 116

QY 61 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAACCTTGACACACAGATGAAGT 120
DB 117 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAACCTTGACACACAGATGAAGT 176
QY 121 CAGTAAGAGGAGGAGGATCAACGCGGCAAAATGNTATCAAGATTACGGATCAAGGTTA 180
DB 177 CAGTAAGAGGAGGAGGATCAACGCGGCAAAATGNTATCAAGATTACGGATCAAGGTTA 236
QY 181 TGTGACCTCTCATGAGAGACCATTAATCAATTAATAATGGCAAGGTTCTCTTATGATGCCAT 240
DB 237 TGTGACCTCTCATGAGAGACCATTAATCAATTAATAATGGCAAGGTTCTCTTATGATGCCAT 296
QY 241 CATCAGTGAAGAGCTCTCATGAAGAGTCCGAATATCAGTTGTAAGAGTTCAGACATTTG 300
DB 297 CATCAGTGAAGAGCTCTCATGAAGAGTCCGAATATCAGTTGTAAGAGTTCAGACATTTG 356
QY 301 CAATGAATCAAGGCTGTTATGTCATTAAGGTAAGCGTAAGTAATACATTAATACCTTAA 360
DB 357 CAATGAATCAAGGCTGTTATGTCATTAAGGTAAGCGTAAGTAATACATTAATACCTTAA 416
QY 361 GGATGACGCTCATGCGGATAATATTCGGAACAAAAGAGAGATTAAACGTCAGAGCAGGA 420
DB 417 GGATGACGCTCATGCGGATAATATTCGGAACAAAAGAGAGATTAAACGTCAGAGCAGGA 476
QY 421 ACGAGTCATTAATCAATCACTCAAGAGC---AGATPATGCTGTTGCTGCAGCCAGAGCCCA 477
DB 477 ACACAGTCATAATCACGGGGGTGTTCTAACGATCAAGCAGTAGTTGCGAGCCAGAGCCCA 536
QY 478 AGCAGCTTATACACGAGATGATGGTATATCTTCAATGCATCTCATATCATTTAGGACAC 537
DB 537 AGCAGCTTATACACGAGATGATGGTATATCTTCAATGCATCTCATATCATTTAGGACAC 596
QY 538 GGTGATGCTTATATCTGCTCACGGCGACCATTAACCATTAATTCCTAAGATGAGTT 597
DB 597 GGTGATGCTTATATCTGCTCACGGCGACCATTAACCATTAATTCCTAAGATGAGTT 656
QY 598 ATCAGTACGAGTATGCTGCTCAGAGAGCCTATTTGGAATGGGAAGCAGGATCTCGTCC 657
DB 657 ATCAGTACGAGTATGCTGCTCAGAGAGCCTATTTGGAATGGGAAGCAGGATCTCGTCC 716
QY 658 TTCTTCAAGTTCTAGTTATTAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCCCAA 717
DB 717 TTCTTCAAGTTCTAGTTATTAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCCCAA 776
QY 718 TCTGACTGTCACCTCCAACTTATCATCAAAATCAGGGGAAACATTTCAAGCCTTTTACG 777
DB 777 TCTGACTGTCACCTCCAACTTATCATCAAAATCAGGGGAAACATTTCAAGCCTTTTACG 836
QY 778 TGAATTTGATGCTTAAACCCCTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTGA 837
DB 837 TGAATTTGATGCTTAAACCCCTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTGA 896
QY 838 CCGAGCGCAATCACAAGTGAACCGGCGAGGTTAGTGTGCTCATGGTAAACCATTA 897
DB 897 CCGAGCGCAATCACAAGTGAACCGGCGAGGTTAGTGTGCTCATGGTAAACCATTA 956
QY 898 CCACTTTATCCCTTATGAACAAATGCTGAATGGAATAAGCAATTTCAAGCCTTTTACG 957
DB 957 CCACTTTATCCCTTATGAACAAATGCTGAATGGAATAAGCAATTTCAAGCCTTTTACG 1016
QY 958 CTTTCTGTTATCGTTCAAAACCATTTGGGTACAGATTCAAGACCCAGAACCAACAGTCCACA 1017
DB 1017 CTTTCTGTTATCGTTCAAAACCATTTGGGTACAGATTCAAGACCCAGAACCAACAGTCCACA 1076
QY 1018 ATCGACTCCGGAACTTAGTCCAAATCGGCAACCTGACCAAAATCTTCAACAGCTCCAAAG 1077
DB 1077 ATCGACTCCGGAACTTAGTCCAAATCGGCAACCTGACCAAAATCTTCAACAGCTCCAAAG 1136
QY 1078 CAAATCAATTTGATGAATAATGTTCAAGAGCTGTTGAAAAGTAGGCGATGTTATGT 1137
DB 1137 CAAATCAATTTGATGAATAATGTTCAAGAGCTGTTGAAAAGTAGGCGATGTTATGT 1196
QY 1138 CTTTGGAGGAGATGGAGTTTCTCGTTATATATCCAGCCAGGATCTTTCAGCAGAAACAGC 1197

DB 1197 CTTTGGAGGAGATGGAGTTTCTCGTTATATATCCAGCCAAAGATCTTTTCAGCAGAAACAGC 1256
QY 1198 AGCAGGCTTATGATAGCAAACTGCCAAGCAGGAAAGTTTATCTCATTAAGCTAGAGCTAA 1257
DB 1257 AGCAGGCTTATGATAGCAAACTGCCAAGCAGGAAAGTTTATCTCATTAAGCTAGAGCTAA 1316
QY 1258 GAAAACCTGACCTCCCATCTAGTATCGAGAAATTTTCAATAAAGCTTATGACTTACTAGC 1317
DB 1317 GAAAACCTGACCTCCCATCTAGTATCGAGAAATTTTCAATAAAGCTTATGACTTACTAGC 1376
QY 1318 AAGAAATCCACCAAGATTTACTTGTATTAATAAGGTTCGAACAAGTTGATTTTGAAGGCTTTGGA 1377
DB 1377 AAGAAATCCACCAAGATTTACTTGTATTAATAAGGTTCGAACAAGTTGATTTTGAAGGCTTTGGA 1436
QY 1378 TAACTGTTTGGAAACGACTCAAGGATGTCNCAAGTGTAAAGTCAAGTTAGTGGANGATAT 1437
DB 1437 TAACTGTTTGGAAACGACTCAAGGATGTCNCAAGTGTAAAGTCAAGTTAGTGGANGATAT 1496
QY 1438 TCTTGGCTTTTGTAGCTCCGATTCGATCCAGAACTTTTAGGAAACCAAAATCGCAAT 1497
DB 1497 TCTTGGCTTTTGTAGCTCCGATTCGATCCAGAACTTTTAGGAAACCAAAATCGCAAT 1556
QY 1498 TACCTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACAACAGAGA 1557
DB 1557 TACCTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACAACAGAGA 1616
QY 1558 CGGTTATATCTTTGTATCTCTGATATAACCAAGTATGAGGGGATGCTTATGTAATCTCC 1617
DB 1617 CGGTTATATCTTTGTATCTCTGATATAACCAAGTATGAGGGGATGCTTATGTAATCTCC 1676
QY 1618 ACATATGACCCATGACCTGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGAGCGC 1677
DB 1677 ACATATGACCCATGACCTGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGAGCGC 1736
QY 1678 AGCCAGGCTTATGCTTAAAGAGAAAGTTTGAACCTCTTCCAGACAGACCATCAGGATTC 1737
DB 1737 AGCCAGGCTTATGCTTAAAGAGAAAGTTTGAACCTCTTCCAGACAGACCATCAGGATTC 1796
QY 1738 AGGAAATCTAGGCAAAAGGAGCAGAGCTATCTACAACCGCTGAAAGCAGCTAAGAA 1797
DB 1797 AGGAAATCTAGGCAAAAGGAGCAGAGCTATCTACAACCGCTGAAAGCAGCTAAGAA 1856
QY 1798 GGTGCACTTGTATGCTTACCTTCAATATATCTGTAGAAAGTCAAAAACGCTAG 1857
DB 1857 GGTGCACTTGTATGCTTACCTTCAATATATCTGTAGAAAGTCAAAAACGCTAG 1916
QY 1858 TTTTATCATCTCATTTATGACCATTAACATCAATTTTGAAGTGTGCTGACCAAGG 1917
DB 1917 TTTTATCATCTCATTTATGACCATTAACATCAATTTTGAAGTGTGCTGACCAAGG 1976
QY 1918 CTTTATGAGCACCCTAAGGGGTATCTCTGAGGATCTTTTGGCGACTGTCAAGTACTA 1977
DB 1977 CTTTATGAGCACCCTAAGGGGTATCTCTGAGGATCTTTTGGCGACTGTCAAGTACTA 2036
QY 1978 TGTGAAATCTCAAAAGAAAGTCCGATTCAGATTAATGTTTTTGGTAAACGCTAGCAGCA 2037
DB 2037 TGTGAAATCTCAAAAGAAAGTCCGATTCAGATTAATGTTTTTGGTAAACGCTAGCAGCA 2096
QY 2038 TGTTCAGAAAGAAACAAAATGTTGCTAGCTGATACCAATCAACCGGAAACCAACAGCAGGA 2097
DB 2097 TGTTCAGAAAGAAACAAAATGTTGCTAGCTGATACCAATCAACCGGAAACCAACAGCAGGA 2156
QY 2098 GAAAACCTCAGACAGAAACCTGAGGAAAG-----AACCC 2133
DB 2157 AGTAAGTACGCAACTCACCTCTGATCTGATGAAGAAAGAAATCAGCTGGTTTAAATCC 2216
QY 2134 TCGAGAGAGAGAAAACGCAAAAGCGGAAACCAAGTCTTCCAAACCAACAGAGAAACCGA 2193
DB 2217 TTCAGCAGATAATCTTTTATAAACCAAGCACTGATACGGAAGAGACAGAGAAAGAGCTGA 2276
QY 2194 AGATCACCAGAGAAATCAGAAAGAACCTCAGGTCCAGACTGAAAAGGTTGAAAGAAAACCT 2253

Db 2277 AGATACACAGATGAGGCTGAATTCCTCAAGTAGAGAAATCTGTATTATAACGCTAAGAT 2336
QY 2254 GAGAGAGCTGAAGATTACTTTGGAAAAATCCAGGAT 2290
Db 2337 AGCAGATGCGAGGCTTGCTTAGAAAAAGTAAACAGAT 2373
RESULT 10
US-10-158-844-94
; Sequence 94, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340P1D1
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 94:
US-10-158-844-94
Query Match 87.3%; Score 1999.2; DB 13; Length 8195;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 2135; Conservative 1; Mismatches 154; Indels 27; Gaps 2;
QY 1 TTGTTCTATGAACCTTGCTGCTACCAAGCTGGTCAAGGTTAAGAAAGAGTCTAATCGAGT 60
Db 5658 TTGTTCTATGAACCTTGCTGCTACCAAGCTGGTCAAGGTTAAGAAAGAGTCTAATCGAGT 5717
QY 61 TTCTTATATAGATGGTATCAGGCTGGTCAAGGCGAAGAACTTGACACCCAGATGAAGT 120
Db 5718 TKCTTATATAGATGGTATCAGGCTGGTCAAGGCGAAGAACTTGACACCCAGATGAAGT 5777
QY 121 CAGTAAGAGGAGGGGATCAACGCGCAACAAATGATCAAGATTACGGATCAAGGTTA 180
Db 5778 CAGTAAGAGGAGGGGATCAACGCGCAACAAATGATCAAGATTACGGATCAAGGTTA 5837
QY 181 TGTGACCTCTCATGAGACCAATTATCACTATATAATGGCAAGGTTCTCTTATGATGCCAT 240
Db 5838 TGTGACCTCTCATGAGACCAATTATCACTATATAATGGCAAGGTTCTCTTATGATGCCAT 5897
QY 241 CATCAGTGAAGAGCTCTCATGAAGAGTCCGAATPATCAGTTGAAGAGATTCAGACATTGT 300
Db 5898 CATCAGTGAAGAGCTCTCATGAAGAGTCCGAATPATCAGTTGAAGAGATTCAGACATTGT 5957

QY 301 CAATGAAATCAAGGGTGGTTATGTCATTAAAGGTAAACCGGTAAATATCTATGTTACCTTAA 360
Db 5958 CAATGAAATCAAGGGTGGTTATGTCATTAAAGGTAAACCGGTAAATATCTATGTTACCTTAA 6017
QY 361 GGATGAGCTCATGCGGATAAATTCGGACAAAAGAGAGATTAAACGTCAGAAAGCAGGA 420
Db 6018 GGATGAGCTCATGCGGATAAATTCGGACAAAAGAGAGATTAAACGTCAGAAAGCAGGA 6077
QY 421 ACGCAGTCATAATCAATACTCAAGAGC--AGATAATGCTGTTGCTGCAGCCAGAGCCCA 477
Db 6078 ACACAGTCATAATCAAGGGGGTGGTCTTAACGATCAAGCAGTAGTTGCGACCGAGGCCCA 6137
QY 478 AGGACGTTATACAAACGAGTATGGGTATATCTTCAATGCACTCTGATATCATTTAGAGACAC 537
Db 6138 AGGACGTTATACAAACGAGTATGGGTATATCTTCAATGCACTCTGATATCATTTAGAGACAC 6197
QY 538 GGGTGATGCTTATATCGTTCTCACGGCGACATTAACCATTAATTCCTTAAGAAATGAGTT 597
Db 6198 GGGTGATGCTTATATCGTTCTCACGGCGACATTAACCATTAATTCCTTAAGAAATGAGTT 6257
QY 598 ATCAGCTAGCGAGTTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGATCTCGTCC 657
Db 6258 ATCAGCTAGCGAGTTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGATCTCGTCC 6317
QY 658 TTCTTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTGACAGAACCAAA 717
Db 6318 TTCTTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTGACAGAACCAAA 6377
QY 718 TGTGACTGTCACCTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACG 777
Db 6378 TGTGACTGTCACCTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACG 6437
QY 778 TGAATGTTATGCTTAACCTTATCAGAACGCGCATGTGGATCTGATGGCCCTTATTTTGA 837
Db 6438 TGAATGTTATGCTTAACCTTATCAGAACGCGCATGTGGATCTGATGGCCCTTATTTTGA 6497
QY 838 CCCAGCGCAAAATCAAGAGTCGAACCCAGAGAGTGTAGCTGCCCTCATGTAACCAATTA 897
Db 6498 CCCAGCGCAAAATCAAGAGTCGAACCCAGAGAGTGTAGCTGCCCTCATGTAACCAATTA 6557
QY 898 CCACCTTATCCCTTATGAACAAATGTCTGAATTTGGAAAAACGAATTTGCTGTATTTTC 957
Db 6558 CCACCTTATCCCTTATGAACAAATGTCTGAATTTGGAAAAACGAATTTGCTGTATTTTC 6617
QY 958 CCTTGTGTTATCGTTCAACCATTTGGGTACAGATTCAGACCCAGAACCAACCAAGTCCACA 1017
Db 6618 CCTTGTGTTATCGTTCAACCATTTGGGTACAGATTCAGACCCAGAACCAACCAAGTCCACA 6677
QY 1018 ATCGACTCCGGAACCTTAGTCCGAACCTCGCAACCTCGCAACCTCCTCAACCCAGCTCCAAG 1077
Db 6678 ATCGACTCCGGAACCTTAGTCCGAACCTCGCAACCTCGCAACCTCCTCAACCCAGCTCCAAG 6737
QY 1078 CAATCCAATTTGATGAAAAATTTGGTCAAGAAAGCTGTTTCGAAAAAGTAGGCGATGTTATGT 1137
Db 6738 CAATCCAATTTGATGAAAAATTTGGTCAAGAAAGCTGTTTCGAAAAAGTAGGCGATGTTATGT 6797
QY 1138 CTTTGAAGGAGATGAGATTTCTCGTTATATATCCAGCCAGGATCTTTTCAGCAGAAAAACAGC 1197
Db 6798 CTTTGAAGGAGATGAGATTTCTCGTTATATATCCAGCCAGGATCTTTTCAGCAGAAAAACAGC 6857
QY 1198 AGCAGGATTTGATAGCAAACTGSCCAAGCAGAGAAAGTTTATCTCATAGCTTAGGAGCTAA 1257
Db 6858 AGCAGGATTTGATAGCAAACTGSCCAAGCAGAGAAAGTTTATCTCATAGCTTAGGAGCTAA 6917
QY 1258 GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1317
Db 6918 GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 6977
QY 1318 AAGAAATTCACCAAGATTTACTTGATTAATTAAGGTCGACAGATTTGATTTGAGGCTTTGGA 1377
Db 6978 AAGAAATTCACCAAGATTTACTTTGATTAATAAGGTCGACAGATTTGATTTGAGGCTTTGGA 7037
QY 1378 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGATAAGTCAAGTCTAGTGGANGATAT 1437

7038 TAACTCTGTGTGAAACGACTCAAGGATGTCTCAAGTGAATAAGTCAAGTTAGTGGATGATAT 7097

1438 TCTTGCCTTCTTTAGTCTCGGATTCGTCAATCCAGAACGTTTATAGGAAACCAAAATGGCGCAAT 1497

7098 TCTTGCCTTCTTTAGTCTCGGATTCGTCAATCCAGAACGTTTATAGGAAACCAAAATGGCGCAAT 7157

1498 TACCTACACTGATGATGAGATTCAGTATGACCAAGTTGGCAGGCAAGTGTACAAACAGAAGA 1557

7158 TACCTACACTGATGATGAGATTCAGTATGACCAAGTTGGCAGGCAAGTGTACAAACAGAAGA 7217

1558 CGGTTATATCTTTTGTGATCCTCGTGATATAACCACTGATGAGGGGATGCTTATGTAACTCC 1617

7218 CGGTTATATCTTTTGTGATCCTCGTGATATAACCACTGATGAGGGGATGCTTATGTAACTCC 7277

1618 ACATATGACCCATAGCCACTGGATTATAAAGATAGCTTTGTCTGAAGCTGAGAGAGCGGC 1677

7278 ACATATGACCCATAGCCACTGGATTATAAAGATAGCTTTGTCTGAAGCTGAGAGAGCGGC 7337

1678 AGCCCAGGCTTATGCTAAAGAGAAAGTTTGACCCCTCCTTCGACAGACCAATCAGAGATTC 1737

7338 AGCCCAGGCTTATGCTAAAGAGAAAGTTTGACCCCTCCTTCGACAGACCAATCAGAGATTC 7397

1738 AGGAATATCTGAGCAAAAGGAGCAGAAGCTATCTACACCGCGTGAAGCAGCTAAAGAA 1797

7398 AGGAATATCTGAGCAAAAGGAGCAGAAGCTATCTACACCGCGTGAAGCAGCTAAAGAA 7457

1798 GGTGCCACTGATCGTATGCGTTTACAATCTTCAATATATCTGTAGAAGTCAAAACGGTAG 1857

7458 GGTGCCACTGATCGTATGCGTTTACAATCTTCAATATATCTGTAGAAGTCAAAACGGTAG 7517

1858 TTTAATCATCTCTCATTTATGACCAATACCAATACATCTCAAAATTTGAGTCGTTTGACGAAG 1917

7518 TTTAATCATCTCTCATTTATGACCAATACCAATACATCTCAAAATTTGAGTCGTTTGACGAAG 7577

1918 CCTTTATGAGGCACCTTAAGGGGTATCTCTTGAGGATCTTTTGGCGCATCTCAAGTACTA 1977

7578 CCTTTATGAGGCACCTTAAGGGGTATCTCTTGAGGATCTTTTGGCGCATCTCAAGTACTA 7637

1978 TGTGCAACATCCAAACGAAAGTCCGCAATTCAGATTAATGGTTTGGTAAACGCTAGCGACCA 2037

7638 TGTGCAACATCCAAACGAAAGTCCGCAATTCAGATTAATGGTTTGGTAAACGCTAGCGACCA 7697

2038 TGTTCAAGRAACAAATAATGGTCAAGCTGATACCAATCAACCGGAATAACCAACGGAGGA 2097

7698 TGTTCGTAAATAATAGGTAGACCAAGCAGTAAACCTGATGAAGATAAGGAACATGATGA 7757

2098 GAAACCTTCAGACAGAAATAACCTCAGGAAGA-----AACCCC 2133

7758 AGTAAGTGTAGCAACTCACCTGATCTGATGAATAAGGAATCAACGCTGGTTTAAATCC 7817

2134 TCGAAGAGAGAAACCGCAAGCGAGAAACCGAGTGTCCAAACCAACAGAGGAACCAAGA 2193

7818 TTCAGCAGATAATCTTTATAAACCAAGCACTGATACGGAAGAGACAGAGGAAGAGCTGA 7877

2194 AGAATCCACGAGGAATCAGAGAACTCTCAGGTCGAGACTGAAAAGGTTGAAGAAAAACT 2253

7878 AGATACCACAGATCAGGCTGAAATTCCTCAAGTAGAGAAATTCGTGTTATTAACCGCTAAGAT 7937

2254 GAGAGAGGCTCAAGATTTATCTTGGAAAAATCCAGAT 2290

7938 AGCAGATCGGAGGCGCTTGTAGAAAAAGTAAACAGAT 7974

SULT 11

-09-884-465A-3

Sequence 3, Application US/09884465A

Publication No. US20030077293A1

PUBLICATION NO: US20070100001
GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Chire Biochem Inc

APPLICANT: Shire Biochem, Inc.

APPLICANT: Hamel, Josee

APPLICANT: Brodeur, Bernard.

835 AGTCAAGCAAGTCAAGTAATGACATTGATAGTCTCTTGAACAGCTCTACAAACCTGCCT 894
797 TTATCAGAACGCCATGTGGAAATCTGATGGCCCTTATTTTCGACCGACGCAAAATCACAAGT 856
895 TTGAGTCBAAGCCATGTAGAAATCTGATGGCCCTTATTTTCGACCGACGCAAAATCACAAGT 954
857 CGAACCGCGAGAGGTGATGTCCTCATCGTAAACCAATACCACTTTATCCCTTATGAA 916
955 CGAACCGCGAGAGGTGATGTCCTCATCGTAAACCAATACCACTTTATCCCTTATGAA 1014
917 CAATGTCGTAATTTGAAAAACGAATTCCTCGTATATTTCCCTTCGTATCGTTCAAC 976
1015 CAATGTCGTAATTTGAAAAACGAATTCCTCGTATATTTCCCTTCGTATCGTTCAAC 1074
977 CATTTGGGTACCAAGTTCAGAACCAAGCAACCAAGTCCACCAATCGACTCCGGAACCTAGT 1036
1075 CATTTGGGTACCAAGTTCAGAACCAAGCAACCAAGTCCACCAATCGACTCCGGAACCTAGT 1134
1037 CCAAGTCGCAACCTGCAACCAAACTCCTCAACAGCTCCAAAGCAATCCCAATGATGAGAAA 1096
1135 CCAAGTCGCAACCTGCAACCAAACTCCTCAACAGCTCCAAAGCAATCCCAATGATGAGAAA 1194
1097 TTGSTCAAGAGCTGTTTCAAAAGTAGGCGATGTTATGTTCTTTTGAAGGAATCGAGTT 1156
1195 TTGSTCAAGAGCTGTTTCAAAAGTAGGCGATGTTATGTTCTTTTGAAGGAATCGAGTT 1254
1157 TCTCGTTATATCCAGCCGACAGATCTTTTCAAGCAAAACAGAGAGGAGGATGATGAGCAA 1216
1255 TCTCGTTATATCCAGCCGACAGATCTTTTCAAGCAAAACAGAGAGGAGGATGATGAGCAA 1314
1217 CTGGCCCAAGCAGGAAGTTTATCTCATAGCTAGGAGCTAAGAAAACTGACCTCCATCT 1276
1315 CTGGCCCAAGCAGGAAGTTTATCTCATAGCTAGGAGCTAAGAAAACTGACCTCCATCT 1374
1277 AGTGATCGAGAAATTTTCAATTAAGCTTATGACTTACTAGCAAGAAATTCACCAAGATTA 1336
1375 AGTGATCGAGAAATTTTCAATTAAGCTTATGACTTACTAGCAAGAAATTCACCAAGATTA 1434
1337 CTTGATTAATAAGTTCGACAGATGATTTTGAAGCTTTGATTAACCTGTTGAAAGACTC 1396
1435 CTTGATTAATAAGTTCGACAGATGATTTTGAAGCTTTGATTAACCTGTTGAAAGACTC 1494
1397 AAGGATGTCNCAAGTGTAAAGTCAAGTGTAGTGGAGATATTTCTGCTCTTACTCG 1456
1495 AAGGATGTCNCAAGTGTAAAGTCAAGTGTAGTGGAGATATTTCTGCTCTTACTCG 1554
1457 ATTGCTCATCCAGAACGTTTGAAGAAACCAATGCGCAAAATTAACCTACACTGATGAG 1516
1555 ATTGCTCATCCAGAACGTTTGAAGAAACCAATGCGCAAAATTAACCTACACTGATGAG 1614
1517 ATTCAAGTAGCCAAATTTGAGGAGGAGTCCATGTAACCTCCACATATGACCATAGCCAC 1576
1615 ATTCAAGTAGCCAAATTTGAGGAGGAGTCCATGTAACCTCCACATATGACCATAGCCAC 1674
1577 CGTGATATAACCTGATGAGGAGGAGTCCATGTAACCTCCACATATGACCATAGCCAC 1636
1675 CGTGATATAACCTGATGAGGAGGAGTCCATGTAACCTCCACATATGACCATAGCCAC 1734
1637 TGGATTAATAAAGATAGTTTCTGTAAGCTGAGAGCGGAGCCAGGCTTATGCTAAA 1696
1735 TGGATTAATAAAGATAGTTTCTGTAAGCTGAGAGCGGAGCCAGGCTTATGCTAAA 1794
1697 GAGAAAGGTTTACCCCTCTTCAGACAGCATCGGATTCAGGAATTAACCTGATGATG 1756
1795 GAGAAAGGTTTACCCCTCTTCAGACAGCATCGGATTCAGGAATTAACCTGATGATG 1854
1757 GGACAGAGCTATCTCAACCGGCTGAAGACGATTAAGAGGTCCTTATGATGATG 1816
1855 GGACAGAGCTATCTCAACCGGCTGAAGACGATTAAGAGGTCCTTATGATGATG 1914
1817 CCTTACCAATCTTCAATATCTGTAGAAGTCAAAACCGTAGTTTAAATCATACCTCAT 1876

1915 CTTTCAATCTTCAATATATCTGTAGAAGTCAAAACCGTAGTTTAAATCATACCTCAT 1974
1877 GACCAATTACCAATAACATCAAAATTTGAGTGGTTTGAAGGAGGCTTTATGAGGACCTAAG 1936
1975 GACCAATTACCAATAACATCAAAATTTGAGTGGTTTGAAGGAGGCTTTATGAGGACCTAAG 2034
1937 GGGTATATCTTCTTGAAGATCTTTTGGCGACTGTCTCAAGTACTATGTCGAACATCCAAACGAA 1996
2035 GGGTATATCTTCTTGAAGATCTTTTGGCGACTGTCTCAAGTACTATGTCGAACATCCAAACGAA 2094
1997 CGTCCGCAATCAGATTAATGTTTGGTGAAGCTAGCGACCATGTTCAAGAAACAAAT 2056
2095 CGTCCGCAATCAGATTAATGTTTGGTGAAGCTAGCGACCATGTTCAAGAAACAAAT 2154
2057 GGTCAAGCTGTATACCAATCAACCGAAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2116
2155 GGTCAAGCTGTATACCAATCAACCGAAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2214
2117 CTTGAGAAAGAAACCCCTCGAGAGAGAAACCCGAAAGCGAGAAACCCAGAGTCTCCAAA 2176
2215 CTTGAGAAAGAAACCCCTCGAGAGAGAAACCCGAAAGCGAGAAACCCAGAGTCTCCAAA 2274
2177 CCAACAGAGAAACCCCTCGAGAGAGAAACCCGAAAGCGAGAAACCCAGAGTCTCCAAA 2233
2275 CCAACAGAGAAACCCCTCGAGAGAGAAACCCGAAAGCGAGAAACCCAGAGTCTCCAAA 2334
2234 GAAAAGGTTGAAGAAAAACTGAGAGAGGCTGAAGATTTTCTTGGAAAAATCCAGAT 2290
2335 GAAAAGGTTGAAGAAAAACTGAGAGAGGCTGAAGATTTTCTTGGAAAAATCCAGAT 2391

RESULT 12

US-09-884-465A-4
; Sequence 4, Application US/09884465A
; Publication No. US2003007793A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-4

Query Match 81.1%; Score 1856.8; DB 10; Length 2647;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 2079; Conservative 0; Mismatches 207; Indels 51; Gaps 5;

QY 2 TGTCTCTATGAACCTTGGTCTGCTCACCAGCTGGTTCAGGTTAAGAAAGTCTAATCCGAGTT 61
DB 102 TGTGCTTATGAACCTTGGTCTGCTCACCAGCTGGTTCAGGTTAAGAAAGTCTAATCCGAGTT 158
QY 62 TCTTATATAGATGGTCTGATCAGGCTGGTCAAAAGCGAGAAACCTTGACACAGATGAAGTC 121
DB 159 TCTTATATAGATGGAAACCAAGCGAGCGCAAAACCGAGAAATTTGACTCTCTGATGAGTT 218
QY 122 AGTAAGAGGAGGAGGATCAACCGCAACAAATNGTNATCAAGATTACGATCAAGGTTAT 181
DB 219 AGCAAGCGTGAAGGAATCAACCGCAACAAATCGTCTCATCAAGATTACGATCAAGGTTAT 278
QY 182 GTGACCTCTCTGAGAGACCATTTATCATTTACTATATGCGAAGGTTCTTATGATGCCATC 241

279 GTGACCTCTCATGGAGACCATATATCATTTACTAATATGGCAAGTCCCTTATGATGCCATC 338
242 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCATCTGAGGATTCAGACATGTC 301
339 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCATCTGAGGATTCAGACATGTC 398
302 AATGAAATCAAGGGTGGTTATGTCTATTAAGGTAAACGGTAAATATCTATGNTACCTTAAG 361
399 AATGAAATCAAGGGTGGTTATGTCTATTAAGGTAAACGGTAAATATCTATGNTACCTTAAG 458
362 GATGAGCTCATGCGGATTAATATTCGGAACAAAGAGAGATTAACGTCAGAGCAGGAA 421
459 GATGAGCTCATGCGGATTAATATTCGGAACAAAGAGAGATTAACGTCAGAGCAGGAA 518
422 CGCAGTCATATCAT-----AACTCAAGAGCAGATAATGCTGTTGCTGAGCCAGA 472
519 CATAGTCAGCATCGTGAAGAGGGACTTCAGCAAAACGATGGTGGGTAGCCCTTTGACCGT 578
473 GCCCAAGGACGTTATACCAACGGATGATGGGTATATCTTCAATGTCATCTGATATCATTTGAG 532
579 TCACAGGAGCGCTACACACAGATGATGGTTATATCTTCAATGCACTGATATCATCGAA 638
533 GACACGGGTGATCTTATATCCTTCTCAGCGGACCATTAACCATTAACATTCCTTAAGAAT 592
639 GATACGGGCGATCTATATCCTTCTCATGGAGATCATTAACCATTAACATTCCTTAAGAAT 698
593 GAGTTATCAGCTAGCAGTACTGCTGACAGAGCCTATTGHAATGGGAAGCAGGATCT 652
699 GAGTTATCAGCTAGCAGTACTGCTGACAGAGCCTTCTCTATCTGCTGGGAAAAATCTG 758
653 CGTCCCT-----CTTCAAGTTCAGTTATAAT 679
759 TCAAAATTAAGAACCTATGCGCGACAAATAGGATTAACATCTCAAGAACAAACTGGGTA 818
680 GCAATCCAGCTCAACCAAGATTGTGACAGAAACCAATCTGACTGTCACTCCAACTTAT 739
819 CTTCTGTAAGCAATCCAGGAACCTACAATATAACACAAAGCAACACGACACTAAC 878
740 CATCA---AATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATGTATGTAACCC 796
879 AGTCAAGCAAGTCAAAAGTAATGACATTTGATAGTCTCTTGAAACAGCTCTACAACTGCCT 938
797 TTATCAGAAAGCCATGTGAATCTGATGGCCTTATTTTCGACCCAGCGCAATCAAGT 856
939 TTGAGTCAAGCCATGTGAATCTGATGGCCTTATTTTCGACCCAGCGCAATCAAGT 998
857 CGAACCGGCAGAGGTAGCTGCTCCTCATGGTAACCAATTACCACTTTATCCCTTATGAA 916
999 CGAACCGGCAGAGGTAGCTGCTCCTCATGGTAACCAATTACCACTTTATCCCTTATGAA 1058
917 CAATGTGTGAATTGGAAAAAGAAATGCTGCTGATTTATTTCCCTTCGTTATCGTTCAAC 976
1059 CAATGTGTGAATTGGAAAAAGAAATGCTGCTGATTTATTTCCCTTCGTTATCGTTCAAC 1118
977 CATTCGGTACAGATTCAGAACCAAGACCAACCAAGTCCCAATCGACTCGGAACTAGT 1036
1119 CATTCGGTACCAAGTTCAGAACCAAGACCAACCAAGTCCCAATCGACTCGGAACTAGT 1178
1037 CCAAGTCCGCAACCTCGACCAATCTCAACAGCTCCCAAGCAATCCAATTGATGAGAA 1096
1179 CCAAGTCCGCAACCTCGACCAATCTCAACAGCTCCCAAGCAATCCAATTGATGAGAA 1238
1097 TTGGTCAAGAGAGCTGTTTCGAAAAAGTAGCGATGGTTATGCTTTGAGGAGAAATGGAGTT 1156
1239 TTGGTCAAGAGAGCTGTTTCGAAAAAGTAGCGATGGTTATGCTTTGAGGAGAAATGGAGTT 1298
1157 TCTCGTTATATCCAGCCAGAGATCTTTGACAGAAACAGCAGCAGGCAATTTGATGACAA 1216
1299 TCTCGTTATATCCAGCCAGAGATCTTTGACAGAAACAGCAGCAGGCAATTTGATGACAA 1358
1217 CTGGCCAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAAACTGACCTCCCATCT 1276

1359 CTGCCAAGCAGGAAAGTTTATCTCATAAGCTAGGAGCTAAGAAAACTGAOCTCCATCT 1418
1277 AGTGATCGAGAAATTTTCAANTAAAGCTTATACCTTACTAGCAAGAAATTCACCAAGATTTA 1336
1419 AGTGATCGAGAAATTTTCAANTAAAGCTTATACCTTACTAGCAAGAAATTCACCAAGATTTA 1478
1337 CTTGATAATAAAGTTCGACAAAGTTGATTTTGAGGCTTTGGATAACTGTTGGAACGACTC 1396
1479 CTTGATAATAAAGTTCGACAAAGTTGATTTTGAGGCTTTGGATAACTGTTGGAACGACTC 1538
1397 AAGGATCTCNCAGTGATAAAGTCAAGTTAGTGGANGATATTTCTTGCCCTTTTACTGCTCG 1456
1539 AAGGATCTCNCAGTGATAAAGTCAAGTTAGTGGANGATATTTCTTGCCCTTTTACTGCTCG 1598
1457 ATTTCGTCATCCAGNAACGTTTAGGAAAACCAATGGGAAAATTTACCTACACTGATGAG 1516
1599 ATTTCGTCATCCAGNAACGTTTAGGAAAACCAATGGGAAAATTTACCTACACTGATGAG 1658
1517 ATTCAAGTAGCAGGTTGGCAGGCAAGTACACAACAGAAAGCGGTTATATCTTTGATCCT 1576
1659 ATTCAAGTAGCAGGTTGGCAGGCAAGTACACAACAGAAAGCGGTTATATCTTTGATCCT 1718
1577 CGTGATAATAACCAAGTAGGAGGGGATGCCCTATGTAACTCCACATATGACCCATAGCCAC 1636
1719 CGTGATAATAACCAAGTAGGAGGGGATGCCCTATGTAACTCCACATATGACCCATAGCCAC 1778
1637 TGAATTAATAAAGTAGTGTCTGAAAGCTGAGAGAGCGGAGCCAGGCTTTATGCTAAA 1696
1779 TGAATTAATAAAGTAGTGTCTGAAAGCTGAGAGAGCGGAGCCAGGCTTTATGCTAAA 1838
1697 GAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGGAAATACCTGAGGCAAAA 1756
1839 GAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGGAAATACCTGAGGCAAAA 1898
1757 GGAGCAGAGCTATCTACAAACCGCTGAAAGCAGCTAAGAAAGTGCCACTTCTGATCGTATG 1816
1899 GGAGCAGAGCTATCTACAAACCGCTGAAAGCAGCTAAGAAAGTGCCACTTCTGATCGTATG 1958
1817 CTTTACAATCTTCAATATATCTGTAGAGAGTCAAAAACGGTAGTTTAAATCATACCTCATTA 1876
1959 CTTTACAATCTTCAATATATCTGTAGAGAGTCAAAAACGGTAGTTTAAATCATACCTCATTA 2018
1877 GACCATTTACCATTAACATCAAAATTTGAGTGGTTTTCGAAAGGCTTTTATGAGGCACTAAG 1936
2019 GACCATTTACCATTAACATCAAAATTTGAGTGGTTTTCGAAAGGCTTTTATGAGGCACTAAG 2078
1937 GGGTATATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGCGAACTCCAAACGAA 1996
2079 GGGTATATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGCGAACTCCAAACGAA 2138
1997 GTCCTCGCATTCAGATAATGGTTTGGTAAACGCTAGCGACCATGTTTCAAAGAAACAAAAAT 2056
2139 GTCCTCGCATTCAGATAATGGTTTGGTAAACGCTAGCGACCATGTTTCAAAGAAACAAAAAT 2198
2057 GGTCAAGCTGATACCAATCAAAACGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2116
2199 GGTCAAGCTGATACCAATCAAAACGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2258
2117 CTTGAGGAGAAAGAACCCCTCGAGAGAGAAACCGCAAAACGAGGAGAAACAGAGTCTCCAAAA 2176
2259 CTTGAGGAGAAAGAACCCCTCGAGAGAGAAACCAAAACGAGGAGAAACAGAGTCTCCAAAA 2318
2177 CCAACAGAGAAACCC---AGAAGATCAACAGAGGAATCAGAGAAACCTCAGGTCGAGACT 2233
2319 CCAACAGAGAAACCCAGAGAGGAATCACCAGAGGAATCAGAGGAACCTCAGGTCGAGACT 2378
2234 GAAAGGTTGAGAGAAACCTGAGAGAGGCTGAGATTTACTTTGAGAAATCCAGAT 2290
2379 GAAAGGTTGAGAGAAACCTGAGAGAGGCTGAGATTTACTTTGAGAAATCCAGAT 2435

RESULT 13

US-10-158-844-243

Sequence 243, Application US/10158844
Publication No. US20040029116A1
GENERAL INFORMATION:
APPLICANT: Kunsch et al.
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude Pentium 3
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10158,844
FILING DATE: 03-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/961,527
FILING DATE: 1997-10-30
APPLICATION NUMBER: US 60/029,960
FILING DATE: 1996-10-31
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB340PID1
INFORMATION FOR SEQ ID NO: 243:
SEQUENCE CHARACTERISTICS:
LENGTH: 2359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 243:
US-10-158-844-243

Query Match 60.1%; Score 1376.6; DB 13; Length 2359;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1410; Conservative 0; Mismatches 14; Indels 18; Gaps 1;

QY 1 TTGTTCTATGAACCTGGTGGTCAACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCCAGT 60
Db TTGTTCTATGAACCTGGTGGTCAACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCCAGT 995
QY 61 TTCTTATATAGATGGTGATCAGGCTGGTCAAAAGCGCAGAAACTTTGACACAGATGAAGT 120
Db TGCTTATATAGATGGTGATCAGGCTGGTCAAAAGCGCAGAAACTTTGACACAGATGAAGT 1055
QY 121 CAGTAAGAGGGGGGATCAACCGCAACCAATNGTATCAAGATTACGGATCAAGTTA 180
Db CAGTAAGAGGGGGGATCAACCGCAACCAATNGTATCAAGATTACGGATCAAGTTA 1115
QY 181 TGTGACCTCTCATGGAGACCAATTATCAATTAATTAATGGAAGGTTCTTATGATGCCAT 240
Db TGTGACCTCTCATGGAGACCAATTATCAATTAATTAATGGAAGGTTCTTATGATGCCAT 1175
QY 241 CATCAGTGAAGAGCTCTCTATGAAGAATCCGAATTATCAGTTGAAGGATTCAGACATGTT 300
Db CATCAGTGAAGAGCTCTCTATGAAGAATCCGAATTATCAGTTGAAGGATTCAGACATGTT 1235
QY 301 CAATGAATCAAGGTTGGTTATGTCATTAAGGTTAAACGGTAAATACATGATGTTACCTTAA 360
Db CAATGAATCAAGGTTGGTTATGTCATTAAGGTTAAACGGTAAATACATGATGTTACCTTAA 1295
QY 361 GGATGAGCTCATGCGGATAATATTCGACAAAGAGAGATTAAACGTCAGAGCAGGA 420
Db GGATGAGCTCATGCGGATAATATTCGACAAAGAGAGATTAAACGTCAGAGCAGGA 1355
QY 421 ACGCAGTCATAATCAATCTCAAGAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGG 480

RESULT 14

US-10-412-862-9
; Sequence 9, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 9
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2451)
; OTHER INFORMATION: n = a, c, t or g

Query Match 43.3%; Score 990.6; DB 13; Length 2451;
Best Local Similarity 67.8%; Pred. No. 1.3e-267;
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

1 TTGTTCTTCTATGAACTGGTCTGCTCACCAGCTGCTCAGGTTTAAAGAGTCTTAATCCAGT 60
57 TTGTTCTTCTACGAGTGGGACTGTATCAAGCTGAACGGTTTAAAGAAA---TAACTCGTGT 113
61 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAAATTCGACACAGATGAAGT 120
114 TTCCTATATAGATGGAATCAACAGCGACGCAAAAGAAACGGAGAAATTGACTCTCTGATGAGT 173
121 CAGTAAAGGAGGGGATCAACGCGGAAACAAATNGTNAACAAGATTACGGATCAAGGTTA 180
174 TAGCAAGCGTGAAGGAATCAATGCTCAGCAAAATCGTCATCAAGATAACAGACCAAGGCTA 233
181 TTGTGACCTCTCATGGAGACCATTAATCAATTAATGCAAGGTTCCCTTATGATGCCAT 240
234 TGCTCACTTCACATGGCGACCATCATCTATTAATCAATGGTAAGGTTCCCTTATGACGCTAT 293
241 CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTAATCAGTTCAAGGATTACAGACATTGT 300
294 CATCAGTGAAGAAATTACTCATGAAAGATCCAACTATAAGCTAAAGATGAGGATATTGT 353
301 CAATGAATCAAGGTTGGTTATGTCATTAAGGTAACGGTAAACGTAATATCTATGNTACCTTAA 360
354 TAATCAGGTTCAAGGTTGGATATGTTTATCAAGGTAGATGGAAAATTAATGTTTACCTTAA 413
361 GGATCAGCTCATGCGGATAATTAATTCGGACAAAGAGAGATTAAGAGTCAAGAGCAGGA 420
414 GGATCCTGCCACGCGGATAACGCTCCGTAACAAAGAGGAAATCAATCGACAAAGAAACAGA 473
421 ACAGAGTCATAATCAT- - - - -AACTCAAGAGCAGATATGCTGTGCTGCAGCCAG 471
474 GCATAGTCAACATCGTAGAGTGGAACTCCAGAAACGATGGTGGCTGTGCTTGGCAAG 533
472 AGCCCAAGAGCCTTATACACCGGATGATGGGTATATCTTCAATGCACTGATATCATTTGA 531
534 TTGCAAGAGCAGCTATCATCAGATGATGTTTATATCTTTAATGCTTCTGATATCATAGA 593
532 GCACAGGTTGATGCTTATATGTTTCTCCTCAGCGGACCATTAACCATTAATCTCTAAGAA 591
594 GGATCTGGTGTGCTTATATGTTTCTCCTCATGAGATCATTAACCATTAATCTCTAAGAA 653
592 TGAGTTATCAGCTAGCGATTAGCTGCTGCAAGGCT- - - - - 629

654 TGAGTTATCAGCTAGCGAGTTGGCTGCTGCAGAGGCTTCCCTATCTGGTCGAGGAAATCT 713
630 - - - - -ATTGGAATGGAAGCAGGAGTCTGGTCTCTTCAAGTTCTAGTTATATAA 678
714 GTCAAAATTCAGAAACCTATTCGCCGACAAATAGCGATAACACATCTCAAGAACAACTGGGT 773
679 TCGAAATCCAGCTCAACCAAGATTGTTCAGAGAACCAATCTGACGTCTCACTCCAACTTA 738
774 ACCTTCTGTAAGCAATCCAGGAACTCAAAATATACTAACCAAGCAACACAGCAACTAA 833
739 TCATCA- - -AAATCAAGGGGAAAAATTTTCAAGSCCTTTTACGTGAAATTTGATCTTAAAC 795
834 CAGTCAAGCAAGTCAAGATTAATGACATTGATAGTCTCTTGAACAGCTCTACAACTGCC 893
796 CTTATCAGAACCCATGTTGAATCTGATGGCCCTTATTTTCGACCCAGCGGAAATCAAG 855
894 TTTGAGTCAACGACATAGTAATCTGATGGCCCTTGTCTTTGATCCAGCAAAATCAAG 953
856 TCGAACCGCAGAGGTGATGCTCCCTCATGTAACCATTAACCTTTATCCCTTATGA 915
954 TCGAACAGCTAGAGTGTGAGTGCACACGAGATCAATTACCACTTATCCCTTACTC 1013
916 ACAATGTCTGAATTTGAAAAACGAATTCCTGATTAATTTCCCTTCGTTTATCGTTCAA 975
1014 TCAATGTCTGAATTTGGAAGAACGAATCGCTGATTAATTTCCCTTCGTTTATCGTTCAA 1073
976 CCATTGGGTACAGATTCAAGACCAAGAACCAAGTCCCAATCGACTCCGGAACCTAG 1035
1074 CCATTGGGTACAGATTCAAGGCCAGAAACCAAGTCCCAAGCTCCGGAACCTAG 1133
1036 TCCAAGTCCGCAACTCTGACCAAAATCCTCAACAGTCCAGCAATCCAAATTTGATGAGAA 1095
1134 TCCAGGCCCGCAACTCTGACCAAAATCTTAAATAGACTCAA- - - - -ATTCTTC 1181
1096 ATTGCTCAAGAAAGCTGTTTGAAGAGTAGGCGATGTTATGCTTTTGAGAGAAATCGAGT 1155
1182 TTTGTTAGTCACTGGTGTACGAAAGTTGGGAGGATGATGTTTGAAGAAAGAGGCAT 1241
1156 TTTCTGTTATATCCAGCAAGGATCTTTTCAGCAAGAACAGCAGCAGGCAATTTGATAGCAA 1215
1242 CTCTCGTTATGCTTTGCAAGAGATTACCATCTGAAACTGTTAAATAATCTTGAAGCAA 1301
1216 ACTGCGCAAGCAGAGAAATTTATCTCATAGCTAGGAGTAAAGAAACGTGACCTCCCATC 1275
1302 GTTATCAAAACAGAGAGTGTTCACACACTTAACTGCTTAAAGAAAGAAATTTGTGCTCC 1361
1276 TAGTGATCGAGAAATTTTCAATAAGGCTTTATGACTTACTAGCAAGAAATTCACCAAGATT 1335
1362 TCGTGACCAAGAAATTTTATGATAAGCAATATACTGTTAACTGAGGCTCATAAAGCCTT 1421
1336 ACTTGATAATAAGGTCGACAAAGTTGATTTTGGAGCTTTGGATAACCTGTTGGAAAGCAT 1395
1422 GTTTGNAATAAAGGTCGTAACTGATTTTCCAAAGCTTTAGACAAATTTATAGAAGCCTT 1481
1396 CAGGATGTCNCAAGTGTAAAGTCAAGTTAGTGAGGATTAATCTTGCTTCTTATAGTCC 1455
1482 GAATGATGAATCGACTAATAAGAAAAATTTGTTAGATGATTTTGGCATTCCTTAGCACC 1541
1456 GATTTCGTCAATCCAGAACGTTTGGAAAAACCAAAATTCGCAAAATTTACCTACATGATGA 1515
1542 AATTACCCATCCAGAGCGACTTGGCAACCAAAATTTCTCAAAATGAGTATATGAAGAGCA 1601
1516 GATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAAAGAGACGGTTATATCTTTTATGATCC 1575
1602 AGTTCTGTTTGTCTCAATTTAGCTGATAGTATACAAAGTCAAGTGGTACATTTTATGA 1661
1576 TCGTGATATAACAGTGTGATGAGGGGATGCTATGTAATCTCCACATATGACCCATAGCAA 1635
1662 ACATGATATAATCAGTGTATGAAGGAGATGATATGTAAGCCTCATATGGGCCATAGTCA 1721
1636 CTGGATTAAAAAGATAGTTTGTCTGAGCTGAGAGACGCGCAGCCAGGCTTATGCTTAA 1695

Db 1722 CTGGATGGAAAGATAGACCTTTCTGATAGGAAAGATTGCGAGCTCAAGCCTATACTAA 1781
QY 1696 AGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTCAGGAATACTCGAGGCAAA 1755
Db 1782 AGAAAAAGGTATCCTACCTCCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGA 1841
QY 1756 AGAGAGAAAGCTATCTTAACCGCGTGAAGAGAGCTTAAGAGGTGCCACTTGATCTGAT 1815
Db 1842 TAGTGCAGCAGCTATTTTACAAATCGTGTGAAGGGGAAAAACGAATTCCTCGTTCGACT 1901
QY 1816 GCTTACAACTCTCAATATATCTGAGAGTCAGAAAGCGGTAGTTTAAATCATACCTCATTA 1875
Db 1902 TCCATATATGTTGAGCATACAGTTGAGGTAAAGCGTAATTTGATTTCTCTATAA 1961
QY 1876 TGACCAATTACATAACATCAAAATTTGAGTGGTTTGACGAAGCGCTTTATGAGGACCTAA 1935
Db 1962 GGATCAATTACCAATAATTAATTTGCTTGGTTTGATGATCACACATACAAAGCTCCAAA 2021
QY 1936 GGGGTATCTCTTGAGGATCTTTTGGGACGTGTCAGTACTATGTCGAACATCCAAAGCA 1995
Db 2022 TGCTATACCTTGGAAGATTTGTTTGGGACGATTAAGTACTACGTAGAACAACCTGACGA 2081
QY 1996 AGCTCCGCACTCAGATTAATGTTTGGTAAAGCTAGCGACCATGTTCAAAGAAAACAAAA 2055
Db 2082 ACCTCCACATCTTAATGATGATGGGCAATGCCAGTGAGCATGTGTTAGGCAAGAAAGA 2141
QY 2056 TGCTCAAGCTGATACCAATCAAAACGGAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2115
Db 2142 CCACAGTGAAGATCCAAATAAGAACTTCAAAGCGGATGAAGAGCCAGTAGAGAAACACC 2201
QY 2116 ACTGAGGAAGAAACCCCTCGAGAAAGA 2144
Db 2202 TGCTGAGCCAGAGTCCCTCAAGTAGAGA 2230

RESULT 15

US-10-412-850-9
; Sequence 9, Application US/10412850
; Publication No. US2004001836A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-686
; CURRENT APPLICATION NUMBER: US/10/412,850
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 9
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2451)
; OTHER INFORMATION: n = a, c, t or g
US-10-412-850-9

Query Match 43.3%; Score 990.6; DB 16; Length 2451;
Best Local Similarity 67.8%; Pred. No. 1.3e-267;
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;
QY 1 TTGTTCTCTGAACTTGCTCGTCCCAAGCTGGTCAGGTAAAGAAAGAGTCTTAATCGAGT 60
Db 57 TTGTTCTTACAGTGTGGACCTGTATCAAGCTAGAACGGTTAAGGAAA---TAATCGGT 113
QY 61 TTCCTATATAGATGGTATCAGGCTGGTCAAAAGGCAAGAAACTTGACACCAGATGAAGT 120

Db 114 TTCCTATATAGATGAAGAACGACGCAAAAACCGAGAATTTGACTCCTGATGAGGT 173
QY 121 CAGTAAGAGGAGGAGGATCAACGCCGAAACAAATNGTNATCAAGATPACGGATCAAGGTTA 180
Db 174 TAGCAAGCGTGAAGGAATCAATGCTGAGCAATCGTCATCAAGATAACAGACCAGGCTA 233
QY 181 TGTGACCTCTCATGGAGACCATATCATATATTAATGSCAAGGTTCCCTTATGATGCCAT 240
Db 234 TGTCACTTCATCATGCGGACCACTATCATTTATCAATGATGTAAGGTTCCCTTATGACGCTAT 293
QY 241 CATCAGTCAAGAGCTCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTGT 300
Db 294 CATCAGTGAAGATTTACTCATGAAGATCCAACTATAAGCTAAAGATGAGGATTTGT 353
QY 301 CAATGAAATCAAGGTTGTTATGTCATTAAAGTAAACGGTAAATACTATGTTNATACCTTAA 360
Db 354 TAATGAGGTCAAGGTTGATATGTTATCAAGGTAGATGGAATACTATGTTTACCTTAA 413
QY 361 GGATCAGCTCATGCGGATAATTTCCGACAAAGAGAGATTAACGTCAGAGACGAGA 420
Db 414 GGATGCTCCCAACGCGGATAACGTCGTAACAAAGAGGAAATCAATCGCAAAAAACAAGA 473
QY 421 ACGCAGTCAATATCAT-----AACTCAAGAGCAGATAATGCTGTGCTGCAGCAG 471
Db 474 GCATAGTCAACATCGTGAAGGTGGAACCTCAAGAAACGATGGTGTGCTGTGCTTGGCAGC 533
QY 472 AGCCCAAGGAGCTTATACAAACGAGTATGTTGATATCTTCAATGCACTCTGATATCATTTGA 531
Db 534 TTCCGCAAGGAGCTATACATGATGATGTTTATATCTTTAATGCTTCTCATATCATAGA 593
QY 532 GGACACGGGTGATGCTTATATGCTTCCTCAGCGGACCAATACCATTACATTCCTTAAGAA 591
Db 594 GGATACTGCTGATGCTTATATGCTTCTCATGGAGATCAATTACCATTACATTCCTTAAGAA 653
QY 592 TGAGTTATCAGCTAGGAGTTAGCTGCTGCAGAGCCT----- 629
Db 654 TGAGTTATCAGCTAGGAGTTGGCTGCTGCAGAGCCTTCTCTATCTGGTCGAGGAAATCT 713
QY 630 -----ATTGGAATGGGAAGCAGGAGTCTCGTCCCTTCTTCAAGTTCTAGTTATATA 678
Db 714 GTCAAATTCAGAAACCTATCGCCGACAAATAGCGATAACACTTCAAGAACAAACTGGGT 773
QY 679 TGCAAATCAGCTCAACCAAGATTTGTGAGAACCAATCTGACTGTCTACTCCAATTAA 738
Db 774 ACCTTCTGTAGCAATCCAGGAACTCAAAATCTATACAAAGCAACACAGCAACACTAA 833
QY 739 TCATCA---AATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATTTGTATGCTAAAC 795
Db 834 CAGTCAAGCAAGTCAAGATTAAGCATTTGATGATCTCTTGAAACAGCTCTACAACTGCC 893
QY 796 CTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTCGACCCAGCGCAATTCACAG 855
Db 894 TTTGAGTCAACGACATGAGATCTGATGSCCTTGTCTTTGATCCAGCACAAATCAACAG 953
QY 856 TCGAACCGGACAGAGTGTAGCTGCTCCCTCATGTTAACCAATTACCATTATTCCTTATGA 915
Db 954 TCGAACAGCTAGAGGTGTGTCAGTGCACACAGGAGATCAATTACCCTTCACTTCTACTC 1013
QY 916 ACAATGCTGAAATTCGAAACAGAAATGCTCGTATTATTCCTTCTGTTATGTTTCAAA 975
Db 1014 TCAATATGCTGAAATGGAAAGCAAGATCGTGTATTTATTTCCCTTCTGTTATGTTTCAA 1073
QY 976 CCATTGGGTACCAGATTCGAAGACCAAGAACCAAGTCCCAATCCAGTCCGGAACCTAG 1035
Db 1074 CATTGGGTACCAGATTCGAAGGCAAGAACCAAGTCCCAACCGACTCCGGAACCTAG 1133
QY 1036 TCCAAAGTCCGCAACCTGCAACCAAAATCCTCAACAGCTCCCAAGCAATCCAAATGATGAGA 1095
Db 1134 TCCAGGCCCGCAACCTGCAACCAAAATCTTAAATAGACTCAA-----ATTCTTC 1181
QY 1096 ATTGTTCAAGAGCTGTTTCGAAAGTAGCGATGTTATGTTTTCGAGGAGATGAGT 1155

1182 TTGCTTGTAGCTGGTACGAAAGTTGGGGAAGGATATGATTTCGAAGAAAGGCGAT 1241
1156 TTCTCGTTATATCCAGCAAGGATCTTTTCACAGAAAACAGCAGCAGGATGATGACAA 1215
1242 CTCTCGTTATGTTTGGGAAGATTTACCATCTGAATCTGTTAAATCTTGAAGCAA 1301
1216 ACTGCCAAGCAGGAAGTTTATCTATAAGCTAGAGCTAGAGAACTGACCTCCATC 1275
1302 GTTATCAAAAACAGAGAGTGTTCACACACTTAACTGCTAAAAAGAAAATGTTGCTCC 1361
1276 TAGTGATCGAGAAATTTCAATAAGCTTATGACTTACTAGCAAGAAATTCACCAAGATT 1335
1362 TCGTGACCAAGAAATTTATGATAAAGCATATATCTGTTAACTGAGGCTCATAAAGCCTT 1421
1336 ACTTGATATAAAGCTCCAGAGTTGATTTTTCAGGCTTTTGGATAAAGCTGTTGAAAGACT 1395
1422 GTTTGNAATAAAGGTCGTAATCTGATTTTCCAGCCTTTAGACAAATTAATTAAGACGCTT 1481
1396 CAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGANGATATTTTGGCTTCTTAGCTCC 1455
1482 GAATGATGAATCGACTAATAAGAAABAATTTGGTAGATGATTTATTGGCATTTCTAGCACC 1541
1456 GATTCGTATCCAGAAAGTTTAGGAAAACCAAAATGCGCAAAATTAACCTACACTGATGATGA 1515
1542 AATTACCATCCAGAGCGACTTGGCAAAACCAAAATTTCTAAATTTGATATACTGAAGACGA 1601
1516 GATTCAGCTAGCAAGTTGGCAGGCAAGTACACAGAAAGCGTTATATCTTTGATCC 1575
1602 AGTTGCTATTTGCTCAATTAGCTGATAAGTATACAGCTCAGATGGTTTACATTTTATGA 1661
1576 TCGTGATATAACAGCTGATGAGGGGATCCCTATGTAATCTCCACATATGACCCATAGCCCA 1635
1662 ACATGATATATCACTGATGAGGAGATGATATGTAACGCTCATATGGGCATAGTCA 1721
1636 CTGGATTAATAAAGATAGTTTCTGTAAGCTGAGAGCGGCGAGCCAGCTTATGCTAA 1695
1722 CTGGATTTGAAAAGATAGCTTTCTGATGAAGAAAAGTTGAGCTCAAGCTTATACTAA 1781
1696 AGAGAAAGTTTACCCCTCTTCGACAGACCATCAGGATTCAGGAATTCAGGAATCTGAGC 1755
1782 AGAAAAGGTTATCTTACCTTCCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGA 1841
1756 AGGACGAGAAGCTATCTCAACCGCTGAAAGCAGCTAAGAGGTCCTCATGCTAT 1815
1842 TAGTGACAGAGCTATTTCAATCTGTTGTAAGGGGAAAACGAAATTCCTGCTGACT 1901
1816 GCCTTACAACTTCAATATCTGTAAGCTCAAAAACGGTAGTTTAACTCATACCTCATTA 1875
1902 TCCATATATGTTGAGCATACAGTTGAGGTTTAAACCGGTAATTTGATTTATCTCATAA 1961
1876 TGACATTAACCATACATCAAAATTTGAGTTGTTGACGAGGCTTTTGTAGGACCTTAA 1935
1962 GGATCATTTACCAATAATTAATAATTTGCTTGGTTGATGATCACACATACAAAGCTCCAA 2021
1936 GGGGTATATCTTTGAGGATCTTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGA 1995
2022 TGGCTATACCTTGGAAAGATTTGTTTGGCAGATTAAGTACTAGTAGAACCCCTGACGA 2081
1996 AGCTCCGATTCAGATAATGTTTGGTAACGCTAGCGACCATGTTTCAAGAAACCAAAA 2055
2082 AGCTCCACATTTAATGATGGATGGGCAATGCCAGTGAGCATGTTGTTAGGCAAGAAAGA 2141
2056 TGGTCAGCTGATACCAATCAACGGAAGAAACCAAGCGAGGAGAAACCTTCAGACAGAAA 2115
2142 CCACAGTGAAGATCCAAATAAGAACTTCAAGCGGATGAGAGCCAGTAGAGGAAACACC 2201
2116 ACCTGAGGAGAAACCCCTCGAAGAGA 2144
2202 TGCTGAGCCAGAAGTCCCTCAAGTAGAGA 2230

RESULT 16

US-10-387-783-9

; Sequence 9, Application US/10387783
; Publication No. US2004005331A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 9
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2451)
; OTHER INFORMATION: n = a, c, t or g
US-10-387-783-9

Query Match 43.38; Score 990.6; DB 16; Length 2451;
Best Local Similarity 67.8%; Pred. No. 1.3e-267;
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;
QY 1 TTGTTCTCTATGAACCTGGTCTGCTACCAAGCTGGTCAAGTTAAGAAAAGAGTCTAAATCGAGT 60
Db 57 TTGTTCTTACAGTTGGGACTGTATCAAGCTAGAACGGTTAAGGAAA---TAATCGTGT 113
QY 61 TTCTTATATAGATGGTGATCAGGCTGGTCAAAAGCGAGAAAACCTTGACACAGATGAAGT 120
Db 114 TTCTTATATAGATGTAAGAAAACAAAGCGACGAAAACCGAGAAATTTGACTCTCTGATGAGT 173
QY 121 CAGTAAGAGGAGGGGATCAACGCCGAACAAATGTTATCAAGATTACGGATCAAGTTTA 180
Db 174 TAGCAAGCGTGAAGGAATCAATGCTGAGCAATCGTCATCAAGATAACAGAACCAAGGCTA 233
QY 181 TGTGACCTCTCATGCGAGACCATTAATCAATCTATAATGCGCAAGGTTCTTATGATGCCAT 240
Db 234 TGTCACTTCATCGCGACCACTATCAATTTACATGTAAGGTTCTTATGACGCTAT 293
QY 241 CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTGT 300
Db 294 CATCAGTGAAGAAATTAATCATGAAGATCCAACTATAAGCTAAAAGATGAGGATATTCT 353
QY 301 CAATGAAATCAAGGGTGGTTATGCTAATTAAGGTAAACGGTAAATCTATGTTACCTTAA 360
Db 354 TAATGAGGTCAAGGGTGGATATGTTATCAAGGTAGATGAAAATACTATGTTTACCTTAA 413
QY 361 GGATCGAGCTCATCGGNTAATATTCGACAAAAGAGAGATTAACGCTCAGAAGCAGA 420
Db 414 GGATGCTGCCCGGATAACGCTCGGTACAAAAGAGGAAATCAATCGACAAAACAAGA 473
QY 421 ACGCAGTCATAATCAT-----AACTCAAGAGCAGATAATGCTGTTGCTGACGCCAG 471
Db 474 GCATAGTCAACATCGTGAAGTGGAACTCCAAAGAAACGATGGTGTGTTGCTTGGCAGCG 533
QY 472 AGCCCAAGGAGGTTATCAACGGATGATGGGTATATCTTCAATGCAATCATGATCATTTGA 531
Db 534 TTCGAAAGGAGCTATACACAGATGATGGTTATATCTTTAATGCTTCTGATATCATAGA 593
QY 532 GGACACGGGTGATGCTTATATCGTTCTCTCAGCGGACCATTAACCATTCATTCCTAAAGAA 591
Db 594 GGATACGGTGAATGCTTATATCGTTCTCTCATGGAGATCATTAACCATTCATTCCTAAAGAA 653
QY 592 TGAGTTATCAGCTAGCGAGTTAGCTGTGCGAAGACCT----- 629

Db 654 TGAGTTATCAGCTAGCGAGTTGGCTGTGTCAGAAAGCCCTTCCTATCTGCTGCGAGAAATCT 713
QY 630 -----ATTGGAATCGGAAGCAGGAGTCTGCTCTTCAAGTTCTTAGTTATAA 678
Db 714 GTCAAAATTCAGAACCTATCGCGCAAAATAGGATACACTTCAGAAACAACCTGGT 773
QY 679 TGCAAAATCCAGCTCAACCAAGATTTGTCAGAGAACCAAAATCTGCTGCTCCTCAACTTA 738
Db 774 ACCTTCTGTAAAGCAATCCAGGAACCTCAAAATCTAAACACAGCAACACAGCACTAA 833
QY 739 TCATCA---AAATCAAGGGGAAACAATTCAGCCCTTTTAGGTAATGATGCTAAACC 795
Db 834 CAGTCAAGCAAGTCAAAAGTAATGACATTTGATAGTCTCTTGAACAGCTCTACAAACTGCC 893
QY 796 CTTATCAGACGCCATGTCGAATCTGATGCTTATTTGACCCAGCCAGCAAAATCAAG 855
Db 894 TTTGAGTCAACGACATGTAAGTCTGATGCTTGTCTTTGATCCAGCAAAATCAAG 953
QY 856 TCGAACCCGCGAGAGGTAGTCTGCTCCTCATGTPAACCAATTTATCCCTTATGA 915
Db 954 TCGAAACAGCTAGAGGTGTCAGTGCACACAGGAGATCAATACCACTTCATCCCTTACTC 1013
QY 916 ACAATGCTGTAATTTGAAAAAGAAATGCTGTAATTTATCCCTCTGTTATCGTTCAA 975
Db 1014 TCAATGCTGTAATTTGAAAGAAAGAAATGCTGTAATTTATCCCTCTGTTATCGTTCAA 1073
QY 976 CAAATGGGTACAGATTCAGAACCAAGACCAACCAAGTCCCAATTCGATCGGACCTAG 1035
Db 1074 CCAATGGGTACCAAGATTCAGAGCCAGAACCAACCAAGTCCCAACCGACTCCGGAACCTAG 1133
QY 1036 TCCAAGTCCGCAACTCCTCACCAATCTCAACAGCTCCAGCAATCAATTTGATGAA 1095
Db 1134 TCCAGGCCCGCAACTCCTCACCAATCTCAATATGACTCA-----ATTCTTC 1181
QY 1096 ATTGGTCAAAAGAGCTTTCGAAAGTAGGCGATGGTATGCTTTGAGGAGATGGAGT 1155
Db 1182 TTTGGTGTAGTCAGCTGTACGAAAGTTGGGAAGGATATGTTTCGAAGAAAGGGCAT 1241
QY 1156 TTTCTGTTATATCCAGCAGAGATCTTTGAGAGAAACAGCAGCAGGCAATGTAGCAA 1215
Db 1242 CTCTCGTTATGCTTTGCGAAGATTTACCATCTGAAACTTTAAATCTTTGAAGCAA 1301
QY 1216 ACTGGCCAGCAGAAAGTTTATCTCAATAGCTAGGAGTAAAGAACTGACCTCCCATC 1275
Db 1302 GTTATCAAAACAGAGAGTGTTCACACACTTTACTGCTAAAAGAAAGAAATGTGTCC 1361
QY 1276 TAGTGATCGAGAAATTTACAAATAGGCTTATGACTTACTAGCAAGAAATTCACCAAGATTT 1335
Db 1362 TCGTGACCAAGAAATTTATGATAAGCATATAATCTGTAACTGAGGCTCATAAAGCCTT 1421
QY 1336 ACTTGATAATAAGGTGCAACAGTTGATTTGAGGCTTTGATTAACCTGTTGGAACGACT 1395
Db 1422 GTTTGNAATAAAGGCTGTAACTGATTTTCAAGCCCTTAGCAAAATTTATAGACGCTT 1481
QY 1396 CAGGATGTCNAGTGTAAAGTCAAGTGTAGTGAGATATCTTGCTCTTTAGCTCC 1455
Db 1482 GAATGATGAATCGACTAATAAGAAATTTGATGATGATTTATTTGCAATCTTCTAGCAC 1541
QY 1456 GATTCGTCATCCAGAGCTTTAGGAAACCAAAATGCGCAAAATTTACTACACTGATGTA 1515
Db 1542 AATTTACCAATCCAGGCACTTGGCAACCAAAATTTCAATTTGATGATGATGAGCA 1601
QY 1516 GATTCAGATAGCAGTTGGCAGCAAGTACACAAAGAGAGCGTTATATCTTTGATCC 1575
Db 1602 AGTTGCTATTCTCAATAGCTGATAAGTATACACAGTCAAGTGTTCATTTTGTATGA 1661
QY 1576 TCGTGATATAACAGTGTAGGGGATGCTTATGTAATCCACATATGACCATAGCA 1635
Db 1662 ACATGATATAATCAGTGTAGGAGATGCAATGTAACGCTCATATGAGGCAATGCA 1721
QY 1636 CTGATTAATAAAGAGATGTTTGTCTGAAGCTGAGAGAGCGGACCCAGGCTTATGCTAA 1695
Db 1722 CTGATTTGAAAGAGTACCTTTCTGATAGGAAAGTTCAGCTCAAGCTTACTATA 1781

QY 1696 AGAGAAAGTTTGACCCCTCTCTCGACAGACCATCAGGATTCAGAAATCTGAGGCAAA 1755
Db 1782 AGAAAAAGGTATCTTACCTCTCATCTCCAGACGCAGATGTAAAGCAAAATCCAACCTGAGA 1841
QY 1756 AGGAGCAGAACCTATCTACAACCGCTGAAGCAGCTGAAGAGGTGCCACTTGTATCGTAT 1815
Db 1842 TAGTGCAGCAGCTATTTACAATCGTGTGAAAGGGGAAACGAATTCCTACTCGTTCGACT 1901
QY 1816 GCCTTACAACTCTCAATATCTAGTAAGTCAAAAACGCTAGTCTTTAATCATACCTCATTA 1875
Db 1902 TCCATATATGTTGAGCATACAGTTGAGTTTAAACCGTAAATTTGATTTCTCTCATAA 1961
QY 1876 TGACCAATTAACCAATCAATCAATTTGAGTGGTTTGAACGAGCCCTTTATGAGCAGCTAA 1935
Db 1962 GGATCAATTAACCAATATTAATTTGCTTGTGTTGATGATCACACATACAAAGCTCCAA 2021
QY 1936 GGGGTATATCTCTGAGGATCTTTTGGCGACTCTCAAGTACTATGTCAGACATCCAAACGA 1995
Db 2022 TGGCTATATCTTGGAGATTTGTTTTCGACGATTAAGTACTAGTAGAACAACCTGACGA 2081
QY 1996 ACCTCCGATTCAGATAATGTTTGGTAACTGCTAGCGACCATGTTCAAAGAAACAAAA 2055
Db 2082 ACCTCCCATCTTAATGATGATGGGCAATGCCAGTGTAGGATGCTTTAGGCAAGAGA 2141
QY 2056 TGGTCAAGCTGATACCAATCAACCGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAA 2115
Db 2142 CCACAGTGAAGATCAATAAGAACTTCAAGCGGATGAAGCCAGTAGAGGAAACACC 2201
QY 2116 ACCTGAGGAAGAACCCCTCGAGAGAGA 2144
Db 2202 TGCTGAGCCAGAGTCCCTCAAGTAGAGA 2230

RESULT 17

US-09-765-272-55
; Sequence 55, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 55:

S-09-765-272-55

Query Match 43.1%; Score 987.6; DB 9; Length 2389;
Best Local Similarity 67.7%; Pred. No. 9.2e-267;
Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

4 TTCCTATGAACTCGTCTGACCAAGCTGGTCAAGTTAAGAAAGAGTCTTAATCGAGTTTC 63
1 TTCTTACAGTTGGAGCTGTATCAAGCTAGAACGTTAAGGAAA--TAAICGTGTTTC 57
64 TTATATAGATGGTGAATCAGCTGGTTCAAAAGCAGAAAACCTTGAACACAGATGAAGTCAG 123
58 CTATATAGATGAAAACAAAGCAGCAGCAAAAACCGAGAATTTGACTCTCTGATGAGGTTAG 117
124 TAAGAGGAGGGATCAACGCGGAACAAATNGTNAATCAAGATTACGGATCAAGTTATGT 183
118 CAAGCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAAGATAACAGACCAAGGCTATGT 177
184 GACCTCTCATGAGACCAATATATCTATCTATATATGCGAAGTTTCTTATGATGCCATCAT 243
178 CACTTCAATGCGGACCACTATCTATTTATTAATGTAAGTTTCTTATGACGCTATCAT 237
244 CAGTGAAGAGCTCCTCATGAAGATCCGAAATTTATCAGTTTGAAGGATTCAGACATTTCAA 303
238 CAGTGAAGATTTACTCATGAAGATCCAACTATTAAGCTAAAAGATGAGATATTTGTAA 297
304 TGAATCAAGGTTGTTATGTCAATTAAGTTAAAGGTAATTAATTAATTAATTAATTAAGGA 363
298 TGAGTCAAGGTTGATATGTTATCAAGCTAGATGAAATTAATTAATTAATTAATTAAGGA 357
364 TGCAGCTCATGCGGATATATTTCCGACAAAGAGAGATTAACAGTCAGAAAGCAGAAAG 423
358 TGCTGCCACGCGGATACGTCGTCGACAAAGAGGAAATCAATCGACAAAACAGAGCA 417
424 CAGTCAATATCAT-----AACTCAAGAGCAGATAATGCTGTGTCGACGCGAGC 474
418 TAGTCAACATCTGAAGTGGAACTCCAGAAACGATGGTGTGCTGGCTTGGCAGCTTC 477
475 CCAAGGACGTTTATCAACGGATGATGGGTATATCTTCAATGGAATCTGATATCATTAAGGA 534
478 GCAAGGACGCTATCTACAGATGATGTTATATCTTTAATGCTTCTGATATCATAGAGGA 537
535 CACGGTGTATCTTATATGTTTCTCAGCGACCATTAACATTAATTTCTTAAGATGA 594
538 TACTGTGTATGCTTATATGTTCTCATGGAGATCATTAACATTTACATTTCTTAAGATGA 597
595 GTTATCAGCTAGCGAGTTAGTCTGTCGAGAGCT----- 629
598 GTTATCAGCTAGCGAGTTGGCTGCTGAGAAAGCTTCTATCTGGTCGAGGAATCTGTC 657
630 -----ATTGGAATGGGAAGCAGGAGATCTCGTCTTCTCAAGTTCTAGTTATATATGC 681
658 AAATTCAGAAACCTATCGCCGCAAAAATAGGATTAACACTTCAAGAACAAATCGGTACC 717
682 AAATCCAGCTCAACCAAGATGTTCAGAGAACCAATCTGACTGTCACTCCAATTTATCA 741
718 TTCTGTAAGCAATCCAGGAATCAAAATACTAAACAAGCAACACAGCAACACTAACAG 777
742 TCA--AAATCAGGGGAAACATTTCAAGCTTTTACGTTGATGCTTAAACCTT 798
778 TCAAGCAAGTCAAGATTAATGACATTTGATGTTCTTTGAAACAGCTCTACAACTGCCTTT 837
799 ATCAGAACGCTATGGAATCTGATGCTTATTTTCGACCCGCGCAAAATCAACAGTCG 858
838 GAGTCAACGACATGTAGAACTGATGCTTCTTTGATCCAGCAAAATCAACAGTCG 897
859 AACCGCCAGAGGTGTAGCTGTCCCTCATGGTAAACATTAACATTTATCCCTTATGAACA 918
898 AACAGCTAGAGGTGTTGCAAGTGGCCACAGGAGATCAATACCACTTCTATCCCTTCTCA 957
919 AATGCTGAAATGGAATAACGAATGCTCGTATTTATTTCCCTTCGTTATCGTTCAAACCA 978

Db 958 AATGCTGAAATGGAAGAACGATCGCTCGTATTTATTTCCCTTCGTTATCGTTCAAACCA 1017
QY 979 TTGGGTAC CAGATTCAAGACCAGAAACAACAGTTCACAAATCGACTCGGAACTAGTCC 1038
Db 1018 TTGGGTAC CAGATTCAAGGCCAGAAACAACAGTTCACAAACCGACTCGGAACTAGTCC 1077
QY 1039 AAGTCCGCAACCTGACCAAAATCTCAACAGCTCCAAAGCAATCCAAATGATGAAAT 1098
Db 1078 AGCCCGCAACCTGACCAAAATCTTAAATAGACTCAA-----ATTCTCTTT 1125
QY 1099 GGTCAAGAGCTGTTTCAAAAGTAGCGATGTTATGTCTTTGAGGAGATGAGTTTC 1158
Db 1126 GGTATGCTAGCTGATCGAAAGTTGGGAGAGGATATGTTTCGAGAAAGGCGATCTC 1185
QY 1159 TCTTATATCCAGCCCAAGGATCTTTTCAGCAGAAACAGCAGCGCATTTGATAGCAAACT 1218
Db 1186 TCGTTATGCTTTGGCGAAGATTTTACCATCTGAAACTGTTAAAAATCTTGAAGCAAGTT 1245
QY 1219 GGCCAAAGCAGAAAGTTTATCTCAATAGCTAGGAGCTAAGAAACTGACCTCCCATCTAG 1278
Db 1246 ATCAAAACAAGAGAGTGTTCACACACTTTAACTGTAAAAAAGAAAAATGTGCTCCCTCG 1305
QY 1279 TGATCGAAGATTTTACAAATAGGCTTATGCTTACTAGCAAGATTTTCAACAGATTTACT 1338
Db 1306 TGACCAAGAAATTTTATGATAAGCATATAATCTGTTAACTGAGGCTCATTAAGCCTTGT 1365
QY 1339 TGATAATAAAGGTCGACAAAGTTGATTTTGGAGCTTTGGATTAACCTGTTGGAAACGACTCAA 1398
Db 1366 TGAAATTAAGGTCGTAATCTGATTTCCAGCCTTAGACAAATTTATTAAGCCTTGAA 1425
QY 1399 GGATGTCNCAAGTGATTAAGTCAAGTTAGTGGAGATATCTTGGCTTCTTAGCTCCGAT 1458
Db 1426 TGATGAATCGACTAATAAAGAAAAATTTGTTAGATGATTTTATTTGGCATTTCTTAGACCAAT 1485
QY 1459 TCGTCATCCAGAAAGTTTAGGAAACCAAAATCGCAAAATTTACTACATGATGATCAGAT 1518
Db 1486 TACCATTCAGAGCGACTTGGCAACCAAAATTTCTCAAAATTTAGTATATCTGAAGACGAGT 1545
QY 1519 TCAAGTAGCCAAAGTTGGCAGGCAAGTACACAAAGAGACGTTATATCTTTGATCCTCG 1578
Db 1546 TCGTATTTGCTCAATTAGCTGATAGTATACACGTCAGATGTTTATTTTGGATGAACA 1605
QY 1579 TGATATAACCAAGTGAAGGGGATGCTATGTAATCTCCACATATGACCCATAGCCACTG 1638
Db 1606 TGATATAATCAGTGAAGGAGATGATATGTAACGCTCATATGGGCCATAGTCACTG 1665
QY 1639 GATTAAAAAAGATAGTTTGTCTGAGCTGAGAGAGCGGCGCCAGGCTTATGCTAAAGA 1698
Db 1666 GATTGGAAAAGATAGCCTTTCTGTATAGGAAAAGTTGAGCTCAAGCCTTATCTAAAGA 1725
QY 1699 GAAAAGTTTGACCCCTCTTCGACAGACCATCAGGATTCAGGAAATPACTGAGGCAAAAGG 1758
Db 1726 AAAAGTATCTTACTCTCATCTCCAGAGCGAGATGTTAAAGCAAAATCCAACTGGAGATAG 1785
QY 1759 AGCAGAACTATCTACAAACCGCTGAAAGCAGCTAAGAGAGTGGCCACTTGTATGCTGTC 1818
Db 1786 TGCAGCAGCTATTTACAAATCGTGTGAAAGGGGAAAAACGAAATTCCACTCGTTCGACTTC 1845
QY 1819 TTACAAATCTCAATATATCTGTAGAAAGTCAAAACCGTAGTTTAAATCATACCTCATATGA 1878
Db 1846 ATATATGTTGAGCATACAGTTGAGGTTAAAAACCGTAAATTTGATTTATTTCTCTCATAGGA 1905
QY 1879 CCATTACCATTAACATAATTTGAGTGGTTTTCAGAAAGCCTTTATGAGCAGCTTAAGGG 1938
Db 1906 TCATTACCATTAATAATTTGATTTGCTTTGATGATCAACATCAAGCTCCAATGG 1965
QY 1939 GTATACCTCTGAGGATCTTTTGGCGACTCTCAAGTACTATGTGCAACATCTCAAAACGACG 1998
Db 1966 CTATACCTTTGAGAGATTTGTTTGGACGATTAAGTACTAGTAGAACACCTCTGACGAAAG 2025
QY 1999 TCCGCAATTCAGATAATGTTTTGGTAACGCTAGCGACCATGTTTCAAGAAAGAAACAAAATGG 2058

Db

2026

TCACATCTTAATGATGGGCAATGCCAGTGAGCATGTGTTAGGCAAGAAAGACCA

2085

QY

2059

TCAAGCTGATACCAATCAACGGAAGAAACCAAGCAGGAGAAACCTCAGACAGAAAGAAC

2118

Db

2086

CAGTGAAGATCCAAATAAGAACTTCAAGCGGATGAGAGCCAGTAGAGGAAACACCTGC

2145

QY

2119

TGAGGAAGAAACCCCTCGAAGAGA

2144

Db

2146

TGAGCCAGAAGTCCCTCAAGTAGAGA

2171

RESULT 18

US-10-158-844-355

; Sequence 355, Application US/10158844

; Publication No. US20040029118A1

GENERAL INFORMATION:

APPLICANT: Kunsch et al.

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R

COMPUTER: Dell Latitude Pentium 3

OPERATING SYSTEM: Windows 98

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/158,844

FILING DATE: 03-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER: US 08/961,527

FILING DATE: 1997-10-30

APPLICATION NUMBER: US 60/029,960

FILING DATE: 1996-10-31

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

INFORMATION FOR SEQ ID NO: 355:

SEQUENCE CHARACTERISTICS:

LENGTH: 973 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 355:

US-10-158-844-355

Query Match 34.7%; Score 794.8; DB 13; Length 973;

Best Local Similarity 98.7%; Pred. No. 1e-212;

Mismatches 810; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY

1

TGTTTCCTATGAACCTGGTCGTCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGT

60

Db

154

TGTTTCCTATGAGCTGGACGTTACCAAGCTGGTCAGGATAGAAAGAGTCTAATCGAGT

213

QY

61

TTCTTTATATAGATGGTATCAGGCTGGTCAAAGGCGAGAAACTTCACACCATGAAGT

120

Db

214

TGCTTATATAGATGGTATCAGGCTGGTCAAAGGCGAGAAACTTCACACCATGAAGT

273

QY

121

CAGTAAGAGGAGGAGGATCAACCCGCAACAAATGATNATCAAGATTACGGATCAAGTTA

180

Db

274

CAGTAAGAGGAGGAGGATCAACCCGCAACAAATGATNATCAAGATTACGGATCAAGTTA

333

QY

181

TGTGACCTCTCATGGAGACCATTAATCAATTAATGAGCAAGGTTCTTATGATGCCAT

240

Db

334

TGTGACCTCTCATGGAGACCATTAATCAATTAATGAGCAAGGTTCTTATGATGCCAT

393

QY

241

CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTTATCAGTTGAAGATTTCAGACATTGT

300

Db

394

CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTTATCAGTTGAAGATTTCAGACATTGT

453

QY

301

CAATGAATCAAGGCTGTTATGCTTAAGGTAACGTAATTAATTAATTAATTAATTAATTA

360

Db

454

CAATGAATCAAGGCTGTTATGCTTAAGGTAACGTAATTAATTAATTAATTAATTAATTA

513

QY

361

GGATGCAGCTCATGGGATTAATTTCCGACAAAGAGATTTAAACGTCAGAAAGCAGGA

420

Db

514

GGATGCAGCTCATGGGATTAATTTCCGACAAAGAGATTTAAACGTCAGAAAGCAGGA

573

QY

421

ACGAGTCTAATATCAATCTCAAGAGCAGATTAATGCTGTGTCAGCCAGAGCCCAAGG

480

Db

574

ACGAGTCTAATATCAATCTCAAGAGCAGATTAATGCTGTGTCAGCCAGAGCCCAAGG

633

QY

481

ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG

540

Db

634

ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG

693

QY

541

TGATGCTTATATGCTTCTCAGCGGACCATTTACCAATTAATCTTAAAGATGAGTTATC

600

Db

694

TGATGCTTATATGCTTCTCAGCGGACCATTTACCAATTAATCTTAAAGATGAGTTATC

753

QY

601

AGCTAGCGAGTTAGCTGCTGAGAGCCCTATTGGAATGGGAAGCAGGATCTCGTCTCTTC

660

Db

754

AGCTAGCGAGTTAGCTGCTGAGAGCCCTATTGGAATGGGAAGCAGGATCTCGTCTCTTC

813

QY

661

TTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCT

720

Db

814

TTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCT

873

QY

721

GACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA

780

Db

874

GACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA

933

QY

781

ATTGATGCTAAACCTTATCAGAACGCTATGGAATCGGATCTG 821

Db

934

ATTGATGCT-AAACCTTATCAGAACGCTATGGAATCGGATCTG 973

RESULT 19

US-10-158-844-258

; Sequence 258, Application US/10158844

; Publication No. US20040029118A1

GENERAL INFORMATION:

APPLICANT: Kunsch et al.

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R

COMPUTER: Dell Latitude Pentium 3

OPERATING SYSTEM: Windows 98

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/158,844

FILING DATE: 03-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER: US 08/961,527

FILING DATE: 1997-10-30

APPLICATION NUMBER: US 60/029,960

FILING DATE: 1996-10-31

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB340PID1

INFORMATION FOR SEQ ID NO: 258:

SEQUENCE CHARACTERISTICS:

LENGTH: 1684 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 258:

US-10-158-844-258

```
Query Match      29.6%; Score 677; DB 13; Length 1684;
Best Local Similarity 99.9%; Pred. No. 2.3e-179;
Matches 688; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1602 ATGCTTATCTACTCACATATGACCCATAGCCACTGGATTAAAGAGATAGTTGCTG 1661
DB 1 ATGCTTATCTACTCACATATGACCCATAGCCACTGGATTAAAGAGATAGTTGCTG 60

1662 AAGCTGAGAGCGGCGAGCCCGCTTATGCTAAAGAGAAAGGTTTGCACCCCTCTCGA 1721
QY
DB 61 AAGCTGAGAGCGGCGA-CCGAGGCTTATGCTAAAGAGAAAGGTTTGCACCCCTCTCGA 119

1722 CAGACCATCAGGATTCAGGAATATCTAGGCAAAAGGAGCAGAGCTATCTCAACGCG 1781
QY
DB 120 CAGACCATCAGGATTCAGGAATATCTAGGCAAAAGGAGCAGAGCTATCTCAACGCG 179

1782 TGAAGCAGCTAAGAGGTGCGACTTATGCTGATGCTTACATCTTCAATATCTAGTAG 1841
QY
DB 180 TGAAGCAGCTAAGAGGTGCGACTTATGCTGATGCTTACATCTTCAATATCTAGTAG 239

1842 AAGTCAAAACGCTAGTTTAAATCATACCTCATATGACCAATACCAATCAATCAAAATTG 1901
QY
DB 240 AAGTCAAAACGCTAGTTTAAATCATACCTCATATGACCAATACCAATCAATCAAAATTG 299

1902 AGTGGTTTCAAGAGCCCTTTATGAGGCACTAAGGGGTATCTCTTTGAGGATCTTTTG 1961
QY
DB 300 AGTGGTTTCAAGAGCCCTTTATGAGGCACTAAGGGGTATCTCTTTGAGGATCTTTTG 359

1962 CGACTGTCAAGTACTATGTGGAACATCCAAACGAAAGTCCGGAATTCAGATAATGTTTG 2021
QY
DB 360 CGACTGTCAAGTACTATGTGGAACATCCAAACGAAAGTCCGGAATTCAGATAATGTTTG 419

2022 GTAAAGCTAGGACCATGTTTCAAGAAACAAATGCTCAAGCTGATACCAATCAAAACGG 2081
QY
DB 420 GTAAAGCTAGGACCATGTTTCAAGAAACAAATGCTCAAGCTGATACCAATCAAAACGG 479

2082 AAAAACCAAGCGAGGAGAAACCTTCAGACAGAAACAAACCTTGAGGAGAAACCCCTCGAGAAG 2141
QY
DB 480 AAAAACCAAGCGAGGAGAAACCTTCAGACAGAAACAAACCTTGAGGAGAAACCCCTCGAGAAG 539

2142 AGAAACCGAAAGCGAGAAACCGAGTCTCCAAACCAACAGAGGAAACCGAGAAATCAC 2201
QY
DB 540 AGAAACCGAAAGCGAGAAACCGAGTCTCCAAACCAACAGAGGAAACCGAGAAATCAC 599

2202 CAGAGAAATCAGAGAACCTCAGCTCAGACTGAAAGGTTGAGAGAAACCTGAGAGAGG 2261
QY
DB 600 CAGAGAAATCAGAGAACCTCAGCTCAGACTGAAAGGTTGAGAGAAACCTGAGAGAGG 659

2262 CTGAGATTTACTTTGGAAAAATCCAGAT 2290
QY
DB 660 CTGAGATTTACTTTGGAAAAATCCAGAT 688
```

RESULT 20

US-09-765-272-181

; Sequence 181, Application US/09765272

; Patent No. US20020061545A1

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

```
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 181:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 181:
US-09-765-272-181
```

Query Match 17.5%; Score 401; DB 9; Length 1342;

Best Local Similarity 65.7%; Pred. No. 1.1e-101;

Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

```
QY 23 CACCAAGCTGCTCAGGTTAAGAAAGAGTCTAATCGAGTTTCTTATATAGATGGTGATCAG 82
DB 8 CAGATCGTTCGAGGAAAATAAGCAATATCGTGTCTTCTTATGTGTCAGCCAG 67

83 GCTGCTCAAAAGGCGAGAAACCTTGACACAGATGAAGTCAAGAGGAGGAGGATCAAC 142
DB 68 TCAAGTCAGAAAGTGAAACCTTGACACAGACAGGTTAGCCAGAAAGAAATTCAG 127

143 GCCGAAACAAATNGVNATCAAGATTCAGGATCAAGGTTATGTGACCTCTCATGGAGCCAT 202
DB 128 GCTGAGCAAAATTGTAATCAAAATTTACAGATCAGGCTATGTAACGTCACCGTGACCAC 187

203 TATCATTTACTATAATGGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCCTCATG 262
DB 188 TATCATTTACTATAATGGGAAAGTTCCTTATGATGCCCTTTTAGTGAAGAACTTTGATG 247

263 AAAGATCCAACTATCACTTGAAGGATTCAGACATTTGCAATGAATCAAGGGTGGTTAT 322
DB 248 AAGATCCAACTATCACTTGAAGGATTCAGACATTTGCAATGAATCAAGGGTGGTTAT 307

323 GTCATTAAAGGTAACCGGTAATATCTATGTTACCTTAAGGATGCAGCTCATGGGATAAT 382
DB 308 ATCATCAAGGTCGATGAAATATTTATGTTCTACCTGAAAGATGCAGCTCATGCTGATAAT 367

383 ATTCGACAAAGAGAGATTAAGCTCAGAGCGAGGAAACGAGTCAATATCAATCACTCA 442
DB 368 GTTCGAACTAAAGATGAATCAATCGTCAAAACAAAGAAC---ATGCAAGATATATGAG 424

443 AGAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGGAGCGTTATACAACGGATATGGG 502
DB 425 AAGGTTAACTTAATGTTGCTGTAGCAAGGTTCTCAGGAGAGATATACGACAAATGATGT 484

503 TATATCTTCAATGCAATCTGATATCAATTTAGGACACGCGGTGATGCTTATATCGTTCCTCAC 562
DB 485 TATGCTTTTAAATCCAGCTGATATTTATCGAAGATACGCGGTAATGCTTATATATCGTTCCTCAT 544
```

563 GGCACCATACCAATCAATTCCTAAGAAATGAGTTATCAGCTAGCGAGTTAGTCTGCGA 622
545 GGAGTCACTATCACTACATCCCAAAAGCGATTTATCTGCTAGTGAAATTAGCAGCAGCT 604
623 GAAGCCTATTGGAATGGGAAGCAGGAGTCTCGTCTCTTCTTCAAGTTCTAGTTAATGCA 682
605 AAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAGTTAGCTATTCTTCAACAGCT 664
683 AATCAGCTCAACCAAGATTGTGTCAGAGAACCAACAATCTGACTCTCTCACTCCAATTTATCAT 742
665 AGTGACAATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAGCCA 712
743 CAAATCAAGGGGAAACCAATTTCAAGCCTTTTACGTGAATTTGATGCTAATAACCTTTATCA 802
713 GCATATAATCTGAAATCTCCAGAGTCTTTGAAGGAATCTTATGATTCACCTAGCGCC 772
803 GAACGCATGTGGAATCTGATGGCCCTTTATTTTCGACCCAGCGCAATATCAAGTCGAAAC 862
773 CAAAGTTACAGTGAATCAGATGGCCTGGTCTTTGACCCCTGCTAAGATTTATCAGTCGTACA 832
863 GCAGAGGTGTAGTCTGCTCATGTGAACCATTAACCTTTATCCCTTATGAACAAATG 922
833 CCAAAATGGAGTTGCGATTCGGCATTCGGCATTCGGCAATTAACCACTTTATTCCTTACAGCAAGCTT 892
923 TCTGAATTCGAAAAACGAATTCGTCTGATTTATTTCCCTTCGT 964
893 TCTGCCTTAGAAGAAAGATTGCAGAAATGGTGGCTATCAGT 934

RESULT 21

US-09-769-787-246
; Sequence 246, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial
; APPLICANT: Gilbert, Christophe
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 246
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-246

Query Match 17.5%; Score 401; DB 10; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.1e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;
23 CACCAAGCTGCTCAGGTTAAGAAAGAGTCTAATCAGTTCTTATATAGATGGTATCAG 82
79 CAGCATCGTTCGCAGGAAATAAGCAACAATAATCGTGTCTTATGTGGATGGCAGCCAG 138
83 GCTGGTCAAAAGGAGAAAATTCGACCATGAGTCAAGTCAAGTCAAGGAGGGGATCAAC 142
139 TCAAGTCAGAAAGTGAATTTGACACACAGCCAGGTTAGCCAGAAAGAGAAATTCAG 198
143 GCCCAAGAAATNGTATCAAGATTACGATCAAGTTATGTCACCTCTCATGGAGACCAT 202
199 GCTGAGCAATTTGATCAAAATTCAGATCAGGCTATGTAACGTCAACGTTGACCCAC 258
203 TATCATTTACTAATGGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGTCTCTCATG 262
259 TATCATTTACTAATGGGAAGTTCTTATGATGCCCTTTTATGTAAGAACTCTTTGATG 318

263 AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGCAATGAAATCAAGGCTGGTTAT 322
319 AAGGATCCAAACTATCAACTTAAAGACGCTGATATTGTCAATGAAGTCAAGGTTGGTTAT 378
323 GTCAATTAAGGTAAGACGCTAAATATCTATGTTTACCTTAAGGATGCGAGTTCATGCGGATAAT 382
379 ATCATCAAGGTCGATGGAATAATTTATGCTACCTGAAAGATGCGAGTTCATGCTGATTAAT 438
383 ATTGGCAAAAAGAGAGATTAAACGTCAGAGAGGAGGACGAGTCATTAATCATTAATCA 442
439 GTTCGAATTAAGATGAATCAATCGTCAAAACAAGAAC-----ATGTCAAAAGATTAATGAG 495
443 AGAGCAGATTAATGCTGTGCTGTCGACGAGACCCAGGACGTTATACAAAGGATGATGGG 502
496 AAGGTTAACTCTAATGTTGCTGTAGCAAGGTTCTCAGGACGAGATATACGCAAAATGATGT 555
503 TATATCTTCAATGATCATCTGATATCAATGAGGACACGCGGTGATGCTTATATCGTTCTCTAC 562
556 TATGTCCTTTAATCCAGCTGATTTATTCGAAGATACGGGTAAATGCTTATATCGTTCTCTCAT 615
563 GGCACCAATTTACCAATTAATCTCTTAAGAAATGAGTTATCAGCTAGCGAGTTAGTCTGTGCA 622
616 GGAGTCACTATCACTACATTTCCCAAAAGCGATTTATCTGCTAGTGAATTAGCAGCAGCT 675
623 GAAGCCTATTGGAATGGGAAGCAGGATCTCGTCTCTTCAAGTCTTAGTTTATTAATGCA 682
676 AAGCAGCATCTGGCTGGAAAAAATATGCAACCGAGTCAGTTAAGCTATTCTTTCAACAGCT 735
683 AATCCAGCTCAACCAAGATTGTCAGAGAACCAACAATCTGACTGTCTCACTCCAACTTATCAT 742
736 AGTGACAATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAGCCA 783
743 CAAATCAAGGGGAAACCAATTTCAAGCCTTTTACGTAATTTGATGCTAATAACCTTATCA 802
784 GCAATAAATCTGAAAATCTCCAGAGTCTTTTGAAGGAATCTATGATTCACCTAGCGCC 843
803 GAACGCCATGTGGAATCTGATGGCCCTTATTTTCGACCCAGCGCAATCAAGTCCGAACC 862
844 CAACGTTACAGTGAATCAGATGGCCTGGTCTTTGACCCCTGCTGAAGATTTATCAGTCGTACA 903
863 GCCAGAGGTGTAGTGTCTCCCATCATGTAACCATTAACCTTTATCCCTTATGAACAAATG 922
904 CCAATGAGGTTGCGAATTCGCGATGGCGACCAATTAACCTTTATTTCTTACAGCAAGCTT 963
923 TCTGAATTTGAAAAACGAATTCGCTGATTTATTTCCCTTCGT 964
964 TCTGCCTTAGAAGAAAGATTGCCAGATGGTGGCTTATCAGT 1005

RESULT 22

US-09-769-744A-23
; Sequence 23, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1455
; TYPE: DNA

ORGANISM: Streptococcus pneumoniae
J-09-769-744A-23

```
Query Match 17.5%; Score 401; DB 10; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.1e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

y 23 CACCAAGCTGCTCAGGTTAAGAAAGAGTCTAATCGAGTCTTCTTATATAGATGCTGATCAG 82
b 79 CAGCATCGTTCCGAGGAAAATAAGCAATAATCGTCTCTCTTATGATGACGCCAG 138
y 83 GCTGGTCAAAAGGCGAAGAACTTACACAGATGAAGTCAAGAGGAGGGGATCAAC 142
b 139 TCAGTTCAGAAAGTGAAGAACTTACACAGATGAAGTCAAGAGGAGGGGATCAAC 198
y 143 GCCGAAACAAATGNGTATCAAGATTAACAGATTAACAGATTAACAGATTAACAGAT 202
b 199 GCTGAGCAAAATGTAATCAAAATTAACAGATTAACAGATTAACAGATTAACAGAT 258
y 203 TATCATTACTATATGCGAGGTTCCCTTATGATGCCATCATCATGAGGATTCCTCATG 262
b 259 TATCATTACTATATGCGAGGTTCCCTTATGATGCCATCATCATGAGGATTCCTCATG 318
y 263 AAAGATCCGAATTAATCAGTTGAAGGATTAACAGATTAACAGATTAACAGATTAACAG 322
b 319 AAGATCCGAATTAATCAGTTGAAGGATTAACAGATTAACAGATTAACAGATTAACAG 378
y 323 GTCATTAAAGGTAACGGTAAATCTATGTTTACCTTAAGGATGCGCTCATGCGGATAT 382
b 379 ATCATCAAGGTCGATGGAATATATTTCTCTGAAAGATGAGCTCATGCTGATAT 438
y 563 GGCGACCAATTAACATTAACATTAACATTAACATTAACATTAACATTAACATTAACAT 622
b 616 GGAGGTCACTATCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 675
y 563 GGCGACCAATTAACATTAACATTAACATTAACATTAACATTAACATTAACATTAACAT 622
b 616 GGAGGTCACTATCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 675
y 623 GAAGCCTATTGGAATGGAAAGAGGATTCGCTCTTCTTCAAGTTCCTGCTGCTGCTGCT 682
b 676 AAGCAGCATCTGGCTGGAAATAATATCAACCGGATCACTTAAGCTATCTTCAACAG 735
y 683 AATCGAGCTCAACAGATTTGTGAGAAACCAATCTGACTGTCTACTCTCACTTATCAT 742
b 736 AGTGACAAATACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCA 783
y 743 CAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTAATTTATGCTTAAACCTTATCA 802
b 784 GCAATTAATCTGAAATCTCCAGAGTCTTTGAGGAATCTTATGATTCACCTAGCGCC 843
y 803 GAAGCGCATGTGGAATCTGATGGCTTATTTTGAACCGGCAATCAAGTCAAGCAAC 862
b 844 CAAGGTTACAGTGAATCAGATGGCTGCTTTGACCCCTGCTAAGATTTATCAGTCTG 903
y 863 GCCAGAGTGTAGCTGTCCTCATGTTGAACATTAACATTTATCCCTTATGAACAAATG 922
b 904 CCAATGAGGTGGATTCGGCATGGGACCAATTAACATTTATCTTCTTACAGCAAGCTT 963
y 923 TCTGAATTTGGAATAACGAATGCTGCTGATTAATTTCCCTGCTTCTGCT 964
b 964 TCTGCCCTTAGAAGAAAGATTTGCCAGAAATGGTGCCTATCAGT 1005
```

RESULT 23

US-10-412-862-7
; Sequence 7, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-862-7

```
Query Match 17.5%; Score 401; DB 13; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.1e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

QY 23 CACCAAGCTGCTCAGGTTAAGAAAGAGTCTAATCGAGTCTTCTTATATAGATGCTGATCAG 82
DB 79 CAGCATCGTTCCGAGGAAAATAAGCAATAATCGTCTCTCTTATGATGACGCCAG 138
QY 83 GCTGGTCAAAAGGCGAAGAACTTACACAGATGAAGTCAAGATTAACAGATTAACAGAT 142
DB 139 TCAGTTCAGAAAGTGAAGAACTTACACAGATGAAGTCAAGATTAACAGATTAACAGAT 198
QY 143 GCCGAAACAAATGNGTATCAAGATTAACAGATTAACAGATTAACAGATTAACAGAT 202
DB 199 GCTGAGCAAAATGTAATCAAAATTAACAGATTAACAGATTAACAGATTAACAGAT 258
QY 203 TATCATTACTATATGCGAGGTTCCCTTATGATGCCATCATCATGAGGATTCCTCATG 262
DB 259 TATCATTACTATATGCGAGGTTCCCTTATGATGCCATCATCATGAGGATTCCTCATG 318
QY 263 AAAGATCCGAATTAATCAGTTGAAGGATTAACAGATTAACAGATTAACAGATTAACAG 322
DB 319 AAGATCCGAATTAATCAGTTGAAGGATTAACAGATTAACAGATTAACAGATTAACAG 378
QY 323 GTCATTAAAGGTAACGGTAAATCTATGTTTACCTTAAGGATGCGCTCATGCGGATAT 382
DB 379 ATCATCAAGGTCGATGGAATATATTTCTCTGAAAGATGAGCTCATGCTGATAT 438
QY 563 GGCGACCAATTAACATTAACATTAACATTAACATTAACATTAACATTAACATTAACAT 622
DB 616 GGAGGTCACTATCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 675
QY 623 GAAGCCTATTGGAATGGAAAGAGGATTCGCTCTTCTTCAAGTTCCTGCTGCTGCTGCT 682
DB 676 AAGCAGCATCTGGCTGGAAATAATATCAACCGGATCACTTAAGCTATCTTCAACAG 735
QY 683 AATCGAGCTCAACAGATTTGTGAGAAACCAATCTGACTGTCTACTCTCACTTATCAT 742
DB 736 AGTGACAAATACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCA 783
QY 743 CAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTAATTTATGCTTAAACCTTATCA 802
DB 784 GCAATTAATCTGAAATCTCCAGAGTCTTTGAGGAATCTTATGATTCACCTAGCGCC 843
QY 803 GAAGCGCATGTGGAATCTGATGGCTTATTTTGAACCGGCAATCAAGTCAAGCAAC 862
DB 844 CAAGGTTACAGTGAATCAGATGGCTGCTTTGACCCCTGCTAAGATTTATCAGTCTG 903
QY 863 GCCAGAGTGTAGCTGTCCTCATGTTGAACATTAACATTTATCCCTTATGAACAAATG 922
DB 904 CCAATGAGGTGGATTCGGCATGGGACCAATTAACATTTATCTTCTTACAGCAAGCTT 963
QY 923 TCTGAATTTGGAATAACGAATGCTGCTGATTAATTTCCCTGCTTCTGCT 964
DB 964 TCTGCCCTTAGAAGAAAGATTTGCCAGAAATGGTGCCTATCAGT 1005
```

Db 736 AGTGCAATTAACGCAATCTGTAGCAA-----AGGATCAACTAGCAAGCCA 783
Qy 743 CAAAATCAAGGGGAAACAACTTCAAGCCTTTTACGTGAATGTATGCTAAACCTTATCA 802
Db 784 GCAAAATAATCTGAAATCTCCAGAGTCTTTTGAAGGAACCTCTATGATTCACCTAGCGCC 843
Qy 803 GAAGCCATGTGAAATCTGATGCGCTTATTTTCGACCGGCAAAATCAAGAGTCGAACC 862
Db 844 CAAAGTACAGTGAATCAGATGCGCTGGTCTTTGACCCCTGTAAGATTAATCAGTCGTACA 903
Qy 863 GCCAGAGGTGTAGTGTCCCTCATGTGTAACCAATPACCACTTTATCCCTTATGAACAAATG 922
Db 904 CCAATGGAGTGGCATTCGCATTCGGATGGCGACCAATPACCACTTTATTCCTTTACAGCAAGCTT 963
Qy 923 TCTGAATGGAAAAACGAATGTGCTGCTATTATTCCTCCCTTCGT 964
Db 964 TCTGCCTTAGAAGAAAAAGATTGCCAGAATGGTGGCTATCACT 1005

RESULT 24
US-10-412-850-7
; Sequence 7, Application US/10412850
; Publication No. US20040001836A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-686
; CURRENT APPLICATION NUMBER: US/10/412,850
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-850-7

Query Match 17.5%; Score 401; DB 16; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.1e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

Qy 23 CACCAAGCTGGTCAGGTTAAGAAAGTCTAATCGAGTTCTTATATAGATGGTATCAG 82
Db 79 CAGCATGTTCCGAGGAAAAATAAGGACAAATATCGTCTCTTATGTGGATGGCAGCAG 138
Qy 83 GCTGGTCAAAAGGCAAGAACTTACACACAGATCAAGTCAAGAGAGGGGGATCAAC 142
Db 139 TCAAGTCAGAAAAAGTGAATACTTACACACAGACAGGTTACCCAGAAAGAGGAATTCAG 198
Qy 143 GCCGAACAAATGTTNATCAAGATTACGGATCAAGGTTATGTGACCTCTCTGAGAGACCAT 202
Db 199 GCTGAGCAAAATGTTAATCAAAATTCAGATCAGGCTATGTAAAGTCAACAGGTGACCAC 258
Qy 203 TATCAATTAATAATGGCAAGGTTCTCTTATGATGCTATCATCATGAGAGTCTCTCATG 262
Db 259 TATCAATTAATAATGGCAAGGTTCTCTTATGATGCTCTTATGATGAGAACTCTGTATG 318
Qy 263 AAAGATCGAATTAATCAATTTGAAGATTCAGACATTTCAATGAATCAAGGGTGGTTAT 322
Db 319 AAGGATCAAACTATCAACTTAAAGACGCTGATTTGTCATTAATGAATCAAGGGTGGTTAT 378
Qy 323 GTCAATTAAGGTAAGCGGTAATAATCTATGTTACCTTAAAGGATGTCAGCTCATCGGATAAT 382
Db 379 ATCATCAAGGTCGATGGAAAAATTTATTTATGTTACTTACCTGAAAGATGTCAGCTCATGCTGATAAT 438

Qy 383 ATTCCGACAAAAGAGAGATTAAACGTCAGAAAGCAGGAACCGCAGTCATAATCATTAATCA 442
Db 439 GTTCGAACATAAGATGAATCAATCGTCAAAACAAGAAC---ATGTCAAAGATATGAG 495
Qy 443 AGACGAGATATGCTGTTGCTGACGCCAGAGCCCAAGAGCGTTATACACAGGATGATGG 502
Db 496 AAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGGACGATATACGCAAAATGATGGT 555
Qy 503 TATATCTTCAATGATCTGATATCATTTGAGGACACGGGTGATGCTTATATCGTTCTCTCAC 562
Db 556 TATGCTTTTAATCCAGCTGATATTATCGAAGATACGGGTAATGCTTATATCGTTCTCAT 615
Qy 563 GGCGACCAATTAACATTTACATTTCTTAAGATGAGTTATCAGTCAGCGAGTGTGCTGCA 622
Db 616 GGAGTCACTATCACTACATTTCCAAAAGCGATTATCTGTAGTGAATTTAGCAGCAGCT 675
Qy 623 GAAGCCTATTGGAATGGGAAGCAGGATCTCGTCTCTTCTTCAAGTCTTAGTTATATGCA 682
Db 676 AAGGACATCTGCTGGGAAAAAATATGCAACCGAGTCAGTTAAGCTATTCTTCAACAGCT 735
Qy 683 AATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTGACTCTCACTCCCACTTATCAT 742
Db 736 AGTGACATAAACACGCAATCTGTAGCAA-----AGATCAACTAGCAAGCCA 783
Qy 743 CAAAATCAAGGGGAAACAACTTTCAAGCCTTTTACGTGAATGTATGCTAAACCTTATCA 802
Db 784 GCAATAAATCTGAAATCTCCAGAGTCTTTTGAAGGAACCTCTATGATTCACCTAGCGCC 843
Qy 803 GAAGCCATGTGGAATCTGATGCGCTTATTTTCGACCGGCAAAATCAAGTCGAACC 862
Db 844 CAAAGTTACAGTGAATCAGATGCGCTGCTTTTGACCTCTGAAGATTATCAGTCGTACA 903
Qy 863 GCCAGAGGTGTAGTGTCTCTCATGTTAACCAATPACCACTTTATCCCTTATGAACAAATG 922
Db 904 CCAATGGAGTTCGATTCGATTCGCGACCAATPACCACTTTATTCCTTTACAGCAAGCTT 963
Qy 923 TCTGAATGGAAAAACGAATGTGCTGCTATTATTCCTCCCTTCGT 964
Db 964 TCTGCCTTAGAAGAAAAAGATTGCCAGAATGGTGGCTATCACT 1005

RESULT 25

US-10-387-783-7
; Sequence 7, Application US/10387783
; Publication No. US2004000531A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-387-783-7

Query Match 17.5%; Score 401; DB 16; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.1e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

Qy 23 CACCAAGCTGGTCAGGTTAAGAAAGTCTAATCGAGTTCTTATATAGATGGTATCAG 82
Db 79 CAGCATGTTCCGAGGAAAAATAAGGACAAATATCGTCTCTTATGTGGATGGCAGCAG 138

; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-09-884-465A-1

Query Match 17.5%; Score 401; DB 10; Length 3120;
Best Local Similarity 65.7%; Pred. No. 1.9e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

QY 23 CACCAAGCTGGTCAAGTTAAGAAAGAGTCTAATCGAGTTTCTTATATAGATGGTGCATCAG 82
Db 79 CAGCATCGTTTCGAGGAGAAATAAGGCAATTAATCGTGTCTTATGTGGTGGCAGCG 138
QY 83 GCTGCTCAAAAGGAGAGAACTTGCACACAGATCAAGTCAAGTAAAGAGGAGGGGATCAAC 142
Db 139 TCAAGTCAGAAAGTGAATCTTGCACACAGATCAAGTAAAGAGGAGGGGATCAAC 198
QY 143 GCCGAACAAATNGTATCAAGATTACCGATCAAGGTTAATGACCTCTCATGGAGACCAT 202
Db 199 GCTGAGCAATTTGTAATCAAAATTTACAGATCAGGCTATGTAACGTCAACCGTGACCAC 258
QY 203 TATCATTTATATATGGCAAGTTCTTATGATGCCATCATCAGTGAAGAGCTCCCTCAG 262
Db 259 TATCATTTATATATGGCAAGTTCTTATGATGCCCTCTTATGAGGAACTCTTTGATG 318
QY 263 AAAGATCGAATTTCAAGTTCAGGATTCAGACATTTGTCATGAATCAAGGGTGGTTAT 322
Db 319 AAGATCCAAACTATCAACTTAAAGACGCTAATTTGTCATGAATCAAGGGTGGTTAT 378
QY 323 GTCAATTAAGGTAAACGGTAAATATCTATGTTACCTTAAAGGATGCGAGTCAATCGGATAAT 382
Db 379 ATCATCAAGTTCGATGGAATATTTATGTCACCTGAAAGATGCGAGTCAATCGTATAAT 438
QY 383 ATTCGGACAAAGAGAGATTAAGCTCAGAGAGAGGAGGAGGATCAATCAATCACTCA 442
Db 439 GTTCGAATCAAGATGAATCAATCGTCAAAACCAAGAAC---ATGTCAAAGATTAATGAG 495
QY 443 AGAGCAGATATGCTGTTGCTGACGAGAGCCCAAGGAGCTTTATCAACGGATGATGGG 502
Db 496 AAGGTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGAGGATATACGACAAATGATGGT 555
QY 503 TATATCTTCAATGCAATCTGATATCAATTGAGGACAGGGGTGATGCTTATATCGTTCTCAC 562
Db 556 TATGCTTTTATCCAGCTGATATATCGAAGATACGGGTAAATGCTTATATCGTTCTCAT 615
QY 563 GCGGACCAATACCAATCAATCTTAAGATGATGTTATCAGTACGAGTGTAGTGTGCA 622
Db 616 GGAGGTCACTATCACTACATTCCTCAAAAGCGATTTATCTGCTAGTGAATTTAGCAGCAGCT 675
QY 623 GAAGCTTATGGAATGGAGGAGGAGGATCTGCTCTTCTTCAAGTTCTAGTTATTAATGCA 682
Db 676 AAAGCAGATCTGGCTGGAATAATATGCAACCGAGTCAAGTTAGCTTATCTTCAACAGCT 735
QY 683 AATCCAGCTCAACCAAGATTTGTCAGAGAACCAATCTGACTGTCTCACTCCAACTTATCAT 742
Db 736 AGTGCAATTAACCGCAATCTGTAGCAA-----AGGATCACTAGCAAGCCA 783
QY 743 CAAATCAAGGGGAGAAACATTTCAAGCCCTTTTACGTAATTTGATGCTAAACCTTATCA 802
Db 784 GCAATTAATCTGAAATCTCCAGAGTCTTTGAAGGAACTCTTATGATTCACCTAGCGCC 843
QY 803 GAACGCCATGTGGAATCTGATGCGCTTTATTTGACCCGAGGCAATCAAGTCCGAACC 862
Db 844 CAACGTTACAGTGAATCAGATGCGCTGCTTTGACCCCTGCTAAGATTATCAGTGTGACA 903

83 GCTGCTCAAAAGGAGAGAACTTGCACACAGATGAAGTCAAGTAAAGAGGAGGGATCAAC 142
b 139 TCAAGTCAGAAAGTGAATCTTGCACACAGATCAAGTAAAGAGGAGGGATCAAC 198
y 143 GCCGAACAAATNGTATCAAGATTACCGATCAAGGTTAATGACCTCTCATGGAGACCAT 202
b 199 GCTGAGCAATTTGTAATCAAAATTTACAGATCAGGCTATGTAACGTCAACCGTGACCAC 258
y 203 TATCATTTATATGCAAGTTCTTATGATGCCATCATCAGTGAAGAGCTCCCTCAG 262
b 259 TATCATTTATATGCAAGTTCTTATGATGCCCTCTTATGAGGAACTCTTTGATG 318
y 263 AAAGATCGAATTTCAAGTTCAGGATTCAGACATTTGTCATGAATCAAGGGTGGTTAT 322
b 319 AAGATCCAAACTATCAACTTAAAGACGCTAATTTGTCACCTGAAAGATGCGAGTCAATCGTATAAT 438
y 323 GTCAATTAAGGTAAACGGTAAATATCTATGTTACCTTAAAGGATGCGAGTCAATCGGATAAT 382
b 379 ATCATCAAGTTCGATGGAATATTTATGTCACCTGAAAGATGCGAGTCAATCGTATAAT 438
y 383 ATTCGGACAAAGAGAGATTAAGCTCAGAGAGAGGAGGATCAATCAATCACTCA 442
b 439 GTTCGAATCAAGATGAATCAATCGTCAAAACCAAGAAC---ATGTCAAAGATTAATGAG 495
y 443 AGAGCAGATATGCTGTTGCTGACGAGAGCCCAAGGAGCTTTATCAACGGATGATGGG 502
b 496 AAGGTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGAGGATATACGACAAATGATGGT 555
y 503 TATATCTTCAATGCAATCTGATATCAATTGAGGACACGGGTGATGCTTATATCGTTCTCAC 562
b 556 TATGCTTTTATCCAGCTGATATATCGAAGATACGGGTAAATGCTTATATCGTTCTCAT 615
y 563 GCGGACCAATACCAATCAATCTTAAGATGATGTTATCAGTACGAGTGTAGTGTGCA 622
b 616 GGAGGTCACTATCACTACATTCCTCAAAAGCGATTTATCTGCTAGTGAATTTAGCAGCAGCT 675
y 623 GAAGCTTATGGAATGGAGGAGGATCTGCTCTTCTTCAAGTTCTAGTTATTAATGCA 682
b 676 AAAGCAGATCTGGCTGGAATAATATGCAACCGAGTCAAGTTAGCTTATCTTCAACAGCT 735
y 683 AATCCAGCTCAACCAAGATTTGTCAGAGAACCAATCTGACTGTCTCACTCCAACTTATCAT 742
b 736 AGTGCAATTAACCGCAATCTGTAGCAA-----AGGATCACTAGCAAGCCA 783
y 743 CAAATCAAGGGGAGAAACATTTCAAGCCCTTTTACGTAATTTGATGCTAAACCTTATCA 802
b 784 GCAATTAATCTGAAATCTCCAGAGTCTTTGAAGGAACTCTTATGATTCACCTAGCGCC 843
y 803 GAACGCCATGTGGAATCTGATGCGCTTTATTTGACCCGAGGCAATCAAGTCCGAACC 862
b 844 CAACGTTACAGTGAATCAGATGCGCTGCTTTGACCCCTGCTAAGATTATCAGTGTGACA 903
y 863 GCGAGGTTAGTATGCTGCTTCAATGTTGATCAATCAACCTTATCCCTTATGAAACAAATG 922
b 904 CCAATAGGTTGCAATTCGCAATGGGACCAATCAACCTTATCCCTTATCCTAGCAGAGCTT 963
y 923 TCTGAAATGGAAAAACGAATTTGCTGCTTATTTATCCCTTCTGCT 964
b 964 TCTGCTTAGAAGAAAGATTTGCAAGATGCTGCTATCATG 1005

RESULT 26
US-09-884-465A-1
; Sequence 1, Application US/09884465A
; Publication No. US2003007293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine

QY	23	CACCNAGCTGGTCAGGTTAAGAAAGAGCTCTAAATCGAGTTCCTTTATATAGATGTGTATCAG	82
Db	1955	CAGCATCGTTCCGCAAGGAATAAAGGACAAATAATCGTGTCTCTTATGTGGATGGCAGCCAG	1914
QY	83	GCTGGTCAAAGGCAGAAAACTTGACACCCAGATGAAGTCAGTAAGAGGGAGGGGATCAAC	142
Db	1915	TCAAGTCAGAAAGTGAAGAACTTGACACCCAGACCAGCTTAGCCAGAAAGAAGAAATTCAG	1974
QY	143	GC GGAA CAAATNGFNATCAAGATTACGGATCAAGGTTATGTGACCTCTCATGAGACCAT	202
Db	1975	GCTGAGCAAAATTTGTAATCAAATAATACAGATCAGGGCTATGTAAACGTACACCGGTGACCAC	2034
QY	203	TATCATTTACTATATATGGCAGGTTCCATTATGATGCCATCATCAGTGAAGAGCTCCTCATG	262
Db	2035	TATCATTTACTATATATGGAAAGTTCCCTTATGATGCCCTCTTTAGTGAAGAACTCTTGATG	2094
QY	263	AAAGATCCGAATTTACGTTGAAGGATTCAGACATTCGTCAATGAAATCAAGGGTGTTAT	322
Db	2095	AAGATCCAAACTATCACTTTAAGACCGCTGATATTGTCAATGAAGTCMAGGGTGTTAT	2154
QY	323	GTCATTTAAGGTTAAACGGTAAATACTATGTNTACCTTAAGGATCAGGTCATTCGGGATAAT	382
Db	2155	ATCATCAAGGTCGATGGAAATATATTGTCTACCTGAAAGATGCAGCTCATGCTGATAAT	2214
QY	383	ATTCGGACAAAAGAAGAGATTAAACGTCAGAGCAGGACCGAGTCATATCATTAATCTCA	442
Db	2215	GTTTGAACATAAGATGAATCAATCGTCAAAAAACAAGAAC --- ATGTCAAAGATTAATGAG	2271
QY	443	AGAGCAGATAATGCTGTGTCGACCCAGAGCCCAAGGACGTTATACAACGGATGATGGG	502
Db	2272	AAGGTTAACTCTAATGTTGCTGTAGCAAGGTTCTCAGGACGATATACGACAAATGATGGT	2331
QY	503	TATATCTTTCAATGCATCTGATATCATTTGAGGACACGGGTGATGCTTATATCGTTCCTCAC	562

Query Match 17.5%; Score 401; DB 13; Length 6867;
Best Local Similarity 65.7%; Pred. No. 3.2e-101;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;
23 CACCAAGCTGGTACAGTAAAGAAAGTCTAATCGAGTTCTTATAGATGGTATGATCAG 82
6707 CAGCATCGTTCGAGGAAATTAAGCAATATCGTGTCTCTTATGTGGATGCGCCAG 6648
83 GCTGTCTAAAGGCGAGAACTTTGACACAGATGAAGTCAAGTAAAGAGGAGGAGGATCAAC 142
6647 TCAAGTCAGAAAGTGAACCTTGACACAGACAGCCAGGTAGCCAGAAAGAGAAATTCAG 6588
143 GCGCAACAAATGTTNATCAAGATTACGATCAAGGTATATGTGACCTCTCATGAGACCAT 202
6587 GCTGAGCAATTTGTAATCAAAATTAACAGATCAGGGCTATGTAACGTACACCGGTGACCAC 6528
203 TATCATTTACTATAATGGCAAGGTTCTTATGATGCGATCATCATGATGAAGAGCTCTCTCATG 262
6527 TATCATTTACTATAATGGCAAGGTTCTTATGATGCGCTCTTTAGTGAAGAACTCTTGATG 6468
263 AAGATCCGAATTTATCAGTTGAGGATTCAGACATTTGTCATGAATCAAGGTGGTGTAT 322
6467 AAGGATCCAAACTATCACTTAAAGACGCTGATTTGTCATGAAGTCAAGGTGGTGTAT 6408
323 GTCAATTAAGGTAAACGCTAAATACTATGTTTAAAGATGAGCTCATGCGGATTAAT 382
6407 ATCATCAAGTTCGATGGAATATTAATGCTCTCTGAAAGATGAGCTCATGCTGATAAT 6348
383 ATTCGCAAAAGAGAGATTAAGTCAAGAGCAGGAAAGCGAGTCAATATCAATCA 442
6347 GTTCGAACCTAAAGATGAATCAATCGTCAAAACAAAGAAC---ATGTCAAAGATTAATGAG 6291
443 AGAGCAGATAATCTCTGTCGAGCAGCAGCCAGGCTTATATCAACGGATGATGGG 502
6290 AAGGTTAACTCTTAATGTTGCTGAGCAAGTCTCAGGAGCATATACGAAATGATGGT 6231
503 TATATCTCAATCACTGATATCAATGAGGACAGCGGCTGATGTTATATCGTTCCCTCAC 562
6230 TATGCTTTAATCCAGCTGATATTCGAGATAGCGGTAATGCTTATATCGTTCCCTCAT 6171
563 GCGCAGCATTACCATTAATTCCTTAAGATGAGTTATCAGTACGAGTGTAGCTGTGCA 622
6170 GGAGGTCATCTACATACATTCCTCAAAAGCGATTTATCTGCTAGTGAATTAGCAGCAGCT 6111
623 GRAGCCTATGGAATGGAGAGCGGATCTCGTCTCTTCAGGTTCTGATTTATATGCA 682
6110 AAGACCATCTGGCTGGAAAAATATGCAACCGAGTCAGTTAAGCTATTTCTTCAACAGCT 6051
683 AATCCAGCTCAACCAAGATTGTGAGAAACCAATCTGACTGTCACTCCCAACTTATCAT 742
6050 AGTGACAATTAACAGGATCTGTAGCAAA-----AGGATCAACTAGCAGGCA 6003
743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATTTGATGCTTAAACCCCTTATCA 802
6002 GCAATTAATCTGAAATCTCCAGAGCTTTTGAAGGAACCTCTATGATTTCACTAGCGCC 5943
803 GAAGCCATGTGGAATCTGATGGCTTATTTTGAACCCAGCGGAATCAAGATCGAAC 862
5942 CAACGTTACAGTGAATCAGATGGCTGCTTTGACCCCTGCTAAGATTTATCAGTCGTACA 5883
863 GCCAGAGGTGAGTGTCCCTCATGGTAAACCATTAACCACTTTATCCCTTATGAACAAATG 922
5882 CCAATGGAGTTGGATTTCCGCATGGGAGGACCATTAACCACTTTATCCCTTACAGCAAGCTT 5823
923 TCTGAAATGGAAAAACGAATTTGCTGATTTATTTCCCTTCGT 964
5822 TCTGCTTAGAAGAAAGATTCGCAATGTTGCTGCTATCAGT 5781

RESULT 29
US-09-884-465A-257
; Sequence 257, Application US/09884465A
; Publication No. US2003007293A1

GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
US-09-884-465A-257

Query Match 16.2%; Score 370; DB 10; Length 819;
Best Local Similarity 73.2%; Pred. No. 4.2e-93;
Matches 534; Conservative 0; Mismatches 130; Indels 66; Gaps 2;
1492 GCAATTAACCTACACTGATGATGAGATTCAAGTACGCCAAGTTGGCAGGCAAGTACACAAC 1551
Db 3 GCAATTAACCTACACTGATGATGAGATTCAAGTACGCCAAGTTGGCAGGCAAGTACACAAC 62
1552 AGAAGACGGTTATATCTTTGATCTCTGATATATACCAAGTATGAGGGGGATGCTTATGT 1611
Db 63 AGAAGACGGTTATATCTTTGATCTAGT----- 90
1612 AACTCCCATATGAGCCCATAGCCATCGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAG 1671
Db 91 -----TGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAG 125
1672 AGCGCAGCCAGGCTTATGCTTAAGAGAAAGTTTGACCCCTCTTCGACAGACCATCA 1731
Db 126 AGCGCAGCCAGGCTTATGCTTAAGAGAAAGTTTGACCCCTCTTCGACAGACCATCA 185
1732 GGATTTCAGAAATCTGAGGCAAAAGGAGCAGAGCTATCTACAACCGGTGAAGCAGC 1791
Db 186 GGATTTCAGAAATCTGAGGCAAAAGGAGCAGAGCTATCTACAACCGGTGAAGCAGC 245
1792 TAAGAAAGTGCATTTGATCGTATGCTTACAAATCTTCAATATCTAGTAAAGTCAAAA 1851
Db 246 TAAGAAAGTGCATTTGATCGTATGCTTACAAATCTTCAATATCTAGTAAAGTCAAAA 305
1852 CGGTAGTTTATCATACCTCATTTATGACCATTAACCATTAACATCAAAATTTGAGTGGTTGA 1911
Db 306 CGGTAGTTTATCATACCTCATTTATGACCATTAACCATTAACATCAAAATTTGAGTGGTTGA 365
1912 CGAAGCGCTTTATGAGGCACTTAAGGGGTATAGTCTTTGAGGATCTTTTGGCGACTGTCAA 1971
Db 366 CGAAGCGCTTTATGAGGCACTTAAGGGGTATAGTCTTTGAGGATCTTTTGGCGACTGTCAA 425
1972 GTACTATGTCGAACATCCAAACGAAACGTCGCGCATTCAGATAATGATGTTTGGTAAAGCTAG 2031
Db 426 GTACTATGTCGAACATCCGGAACGCTAGT-----GACCATGTTTGGTAAAAATAAGGC 476
2032 CGACCATGTTCAAGNAACAAAATGTTCAAGCTGATACCAATCAACCGAAAAACCAAG 2091
Db 477 AGACCAAGATAGTAAACCTGATGAAGATGAAGAACATGATGAAGTGAAGTGAAGCAACTCA 536
2092 CGAGGAGAAACCTCAGACAGAAACCTGAGGAAGAAACCCCTCGAGAGAGAAACCGCA 2151
Db 537 CCTGAACTGATGAAGAAAGAGATCACGCTGGTTTAAATCTTTCAGCAGATATCTTTA 596
2152 AAGCGGAGAAACAGAGTCTTCCAAAACCAACAGAGGAACCAAGAGAAATCACACGAGGAATC 2211

Db 597 TAAACCAAGCACTGATACGGAAGACACAGAGAAAGTGAAGATACCAAGATGAGGC 656

Qy 2212 AGAAGAACCT 2221

Db 657 TGAATTCCT 666

```

RESULT 30
US-10-324-143-10
; Sequence 10, Application US/10324143
; Publication No. US20030232976A1
; GENERAL INFORMATION:
; APPLICANT: HAMEL, JOSEF
; APPLICANT: CHARLAND, NATHALIE
; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: MARTIN, DENIS
; APPLICANT: BLAIS, NORMAND
; APPLICANT: OVELETTE, CATHERINE
; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 55190-56
; CURRENT APPLICATION NUMBER: US/10/324,143
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/341,252
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artific
; OTHER INFORMATION: modified Streptococcus
US-10-324-143-10

```

Query Match 16.2%; Score 370; DB 16; Length 819;
Best Local Similarity 73.2%; Pred. No. 4.2e-93;
Matches 534; Conservative 0; Mismatches 130; Indels 66; Gaps 2;

[illegible]

QY		1552	AGAAGACGGTTATATCTTGTGATTAACCAAGTGATGAGGGGGATGCCTTAGT	1611
Dh		63	AGAAGACGGTTATATCTTGTGATTAACCAAGTGATGAGGGGGATGCCTTAGT	90

QY	1612	ARCTCCACATATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAG	1671
nb	91	-----TGCATTTAAAAAGATAGTTTGTCTGAAGCTGAGAG	125

QY	1672	AGCGGCAGCCCGAGCTTATGCTAAAGAGAAAAGCTTTGACCCCTCTCTTCGACAGACCATCA	1731
ph	176	AGCGGCAGCCCGAGCTTATGCTAAAGAGAAAAGCTTTGACCCCTCTCTTCGACAGACCATCA	185

QY	1732	GGATTTCAGGAATACTGAGGCAAAAGGAGCAGAAAGCTATCTCAACCCCGTGAAGCAGC	1791
nb	186	CGATTTCAGGAATACTGAGGCAAAAGGAGCAGAAAGCTATCTCAACCCCGTGAAGCAGC	245

QY	1792	TAAGAAAGGTGCCACTTGATCGTATGCCTTACAATCTTCAATATAC	1851
ST	246	TATATACCTTCGCACTTCATCGCTATCGCCCTTACCTCTTCGTATC	305

[illegible]

QY	1912	1971
QY	CGAAGGCCTTTATGAGGCACCTAAGGGGTATACCTTCGAGGATCTTTTGGCGACTGTCAA	1971
ST	CGATTCGGCGCTTTTATGAGGCACCTAAGGGGTATACCTTCGAGGATCTTTTGGCGACTGTCAA	495

QY 1972 GTACTATGTCGAACATCCAAACGAACGTCGCGCATTGCTTTGGTAAACGCTAG 2031

Db	426	GTACTATGTCGAACCGGGAACGCTAGT-----GACCATGTTTCGTAATAATAAGGC	476
Qy	2032	CGACCCATGTTCAAAGAAAACAAAATGGTCTAGCTGATACCAATCAACCGAGAAAAACCAAG	2091
Db	477	AGACCAAGATAGTAAACCTGATGAAGATAGGAACATGATGAAGTAGTAGTGAGCCAACTCA	536
Qy	2092	CGAGGGAACCTTCAGACAGAAAACCTGAGGAAGAAACCCCTCGAGAAGAGAAAACCGCA	2151
Db	537	CCCTGCAATCTGATGAAGAAAGAGATCACGCTGTTTAAATCCTTCAGCAGATAACTCTTTA	596
Qy	2152	AAGCGAAGAACCCAGAGTCTCAAAAACCAACAGAGGAACCGAAGAGATTCACCGAGGAATC	2211
Db	597	TAAACCAAGCACTGTATACGGAAGAGACAGAGGAGAGAGCTGAAGATACCAACATGAGGC	656
Qy	2212	AGAAGAACCT	2221
Db	657	TGAAATTCTCT	666

```

RESULT 31
US-10-324-143-4
; Sequence 4, Application US/10324143
; Publication No. US20030232976A1
; GENERAL INFORMATION:
; APPLICANT: HAMEL, JOSEE
; APPLICANT: CHARLAND, NATHALIE
; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: MARTIN, DENIS
; APPLICANT: BLAIS, NORMAND
; APPLICANT: OUELETTE, CATHERINE
; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 55190-56
; CURRENT APPLICATION NUMBER: US/10/324,143
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/341,252
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence
; OTHER INFORMATION: modified Streptococcus
; OTHER INFORMATION: sequence
US-10-324-143-4

```

	Query Match	16.1%	Score 368.4;	DB 16;	Length 816;
	Best Local Similarity	73.0%	Freq. No. 1.2e-94;		
	Matches 533;	Conservative	0;	Mismatches 131;	Indels 66; Gaps 2;
QY	1492	GCAAAATTACCTACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGCGCAAGTACACAAC	1551		
Db	3	GCAAAATTACCTACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGCGCAAGTACACAAC	62		
QY	1552	AGAAGACGGTTATATCTTTTGATCTCTGGTATATACCAAGTCAGTGTATGAGGGGGATGCTATGT	1611		
Db	63	AGAAGACGGTTATATCTTTTGATCTCTGGTATATACCAAGTCAGTGTATGAGGGGGATGCTATGT	90		

QY	1612	AACCTCCATATGACCCCATAGCCACTCGATTAAAAAAGATAGTTTGTCTGAAGCTGAGAG	1671
Dd	91	-----TGGAATTAAAAAGATAGTTTGTCTGAAGCTGAGAG	125
QY	1672	AGCGCAGCCCAGGGTTATGCTTAAGAGAAAAGTTTGACCCCTCCTTCGACAGACCATCA	1731
Dd	126	AGCGCAGCCCAGGGCTTATGCTTAAGAGAAAAGTTTGACCCCTCCTTCGACAGAACCA	185
QY	1732	GGATTACGAAATACTGAGGCARAAAGNGCAGAAGCTATCTACAACCGCGTGAAGCAGC	1791
Dd	186	GGATTACGAAATACTGAGGCARAAAGNGCAGAAGCTATCTACAACCGCGTGAAGCAGC	245
QY	1792	TAAAGAAGGTGCCACTTTGATCGTATGCCCTTACAATCTTCAATATATCTGTAGAAGTCAAAA	1851

Query Match	15.9%;	Score 363.6;	DB 16;	Length 2721;
Best Local Similarity	72.6%;	Pred. No. 6e-91;		
Matches 530; Conservative	0;	Mismatches 134;	Indels 66;	Gaps 2;
1492 GC A A A T T A C C T A C A C T G A T G A T G A G A T T C A A G T A G C C A A G T T G G C A G G C A A G T A C A C A C 1551				
3 GC A A A T T A C C T A C A C T G A T G A G A T T C A G T A G C C A A G T T G G C A G G C A A G T A C A C A C 62				
1552 A G A A G A C G G T T A T A T C T T T G A T C C T C G T G A T A T A A C C A G T G A T A G G G G A T G C C T A T G T 1611				
63 A G A A G A C G G T T A T A T C T T T G A T G G T G T T ----- 90				
1612 A A C T C C A C A T A T G A C C C A P A G C C A C T G G A T T A A A A A G A T A G T T G T C T G A A C T G A G A G 1671				
91 ----- T G G A T T A A A A A A G A T A G T T G T C T G A A C T G A G A G 125				
1672 A G C G G C A G C C A G G C T T A T G C T T A A A G A G A A A G T T T G A C C C C T C C T T C G A C A C A C A T C A 1731				
126 A G C G G C A G C C A G G C T T A T G C T T A A A G A G A A A G T T T G A C C C C T C C T T C G A C A C A C A C C A 185				
1732 G G A T T C A G G A A A T A C T G A G C C A A A A G A G C A G A A G C T A T C T A C A A C C G C G T G A A A G C A G C 1791				
186 G G A T T C A G G A A A T A C T G A G C C A A A A G A G C A G A A G C T A T C T A C A A C C G C G T G A A A G C A G C 245				
1792 T A A G A A G G T G C C A C T T G A T C G T A T G C C T T A C A A T T C T A A T A T A C T G T A A A G T C A A A A 1851				
246 T A A G A A G G T G C C A C T T G A T C G T A T G C C T T A C A A T C T T C A G T A C T A C T G A A G T C A A A A 305				
1852 C G G T A G T T T A A T C A T A C T C A T T A T A C C A T T A C C A T A C A A T C A A A T T T C A G T G G T T T G A 1911				
306 C G G T A G T T T T A A T C A T A C C T C A T C A T A C C A T T A C C A T A A C A A T T T C A G T G G T T T G A 365				
1912 C G A A G C C T T T A T A G A G C C A C T A A G G G T A T A C T C T T G A G A T C T T T T G C G C A C T G T C A A 1971				
366 C G A A G C C T T T A T A G A G C C A C T T A A G G G T A T A G T C T T T G A G A T C T T T T G C G C A C T G T C A A 425				
1972 G T A C T A T G T C G A A C A T C C A A A C G A A C G T C G C A T T C A G A T A A T G G T T T T G G T A A C G C T A G 2031				
426 G T A C T A T G T C G A A G G G G G A A C C T A G T ----- G A C C A T G T T C G T A A A A A T A A A G G C 476				
2032 C G A C C A T G T T C A A A G A A C A A A A A T G G T C A G C T G A T A C C A T C A A A C G A A A A A C C A A G 2091				
477 A G A C C A A G A T A G T A A A C C T G A T A A G A T A A G G A A C A T G A T G A A G T A A G T G A G C C A C T C A 536				
2092 C G A G G A G A A A C C T C A C A G A G A A A A A C T G A G G A A G A A A C C C C T C G A A A G A G A A A C C G C A 2151				
537 C C C T G A A T C T G A T G A A A A A G A G A A T C A C G C T G G T T T A A A T C C T T C A G C A G A T A T C T T T A 596				
2152 A A G C G A G A A A C C A G A G T C T C A A A A C C A A C A G A G G A A C C A G A A G A A T C A C C A G A G A A T C 2211				
597 T A A A C C A A G C A C T G A T A C G A A G A G A C A G A G G A A A A G C T G A A G A T A C C A C A G A T A G A G C 656				
2212 A G A A G A A C C T 2221				
657 T G A A A T T C C T 666				

```

RESULT 34
US - 09-884-465A-9
/ Sequence 9, Application US/09884465A
/ Publication No. US20030077293A1
/ GENERAL INFORMATION:
/ APPLICANT: Shire Biochem, Inc.
/ APPLICANT: Hamel, Josee
/ APPLICANT: Brodeur, Bernard
/ APPLICANT: Martin, Denis
/ APPLICANT: Charland, Nathalie
/ APPLICANT: Ouellet, Catherine
/ TITLE OF INVENTION: Streptococcus A
/ FILE REFERENCE: 055190-0044
/ CURRENT APPLICATION NUMBER: US/09/8
/ CURRENT FILING DATE: 2001-06-20

```

```

; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2528
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-9

```

Query Match.	15.1%;	Score 346.6;	DB 10;	Length 2528;
Best Local Similarity	70.4%;	Pred. No. 3.5e-86;		
Matches 476;	Conservative 0;	Mismatches 197;	Indels 3;	Gaps 1;
QY	23	CACCAAGCTGGTCAGGTTAAAGAAAGAGCTCAATCGAGTTCTTATATAGATGGTGTATCAG	82	
Db	19	CAGCATCGTTGCGACGGAATATAGGAACAATAUGGTGTCTTATGTGATGGCAGCCAG	78	
QY	83	GCTGGTCAAAAGGCGAAGAACTTGACACCCAGATGCAATCAGTAAAGAGGGAGGGATCAAC	142	
Db	79	TCAAGTCAGAAAAAGTGAAAACTTGACACCCAGACCCAGGTTAGCCAGAAAGAGGAATTCAG	138	
QY	143	GCAGAACAAATNGTNATCAAGATTACCGATCAAGGTTATGTGACCTCTCATGGAGACCAT	202	
Db	139	GCTGAGCAAAATTTGTAATCAAAATTTACAGATCAGGGCTATGTACGTCAACGGTGATCAC	198	
QY	203	TATCATTACTATTAATGCAAGGTTCCCTTATGATGCCATCATCAGTGAAGAGCTCCTCATG	262	
Db	199	TATCATTTACTATTAATGGGAAAGTTCCCTTATGATGCCCTCTTTAGTGAAGAACTCTTGATG	258	
QY	263	AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGTCAATGAAATCAAGGGTGGTTAT	322	
Db	259	AAAGATCCAAACTATCAACTTAAAGACGCTGATATTGTCAATGAAGTCAAGGGTGGTTAT	318	
QY	323	GTCAATTAAGGTAAACGGTAAATCATATGTNTWTAACCTTAAGGATGACGCTCATGCGGATTAAT	382	
Db	319	ATCATCAAGGTCGATGCGAAAAATATATGTCTTACCTGAAAGATGACGCTCATGCTGATAAT	378	
QY	383	ATTCCGACAAAAGAGAGATTTAAACGTCAGAGCAGGAACGAGTCAATAATCATAACTCA	442	
Db	379	GTTCCGAATTAAGATGAATCAATCGTCAAAAACAAGACA---TGTCAAAGATTAATGAG	435	
QY	443	AGACGACATAATGCTGTTGCTGACGACGAGCCCAAGGACGTTATACACGGATGATCGG	502	
Db	436	AAAGTTAACTTAATGTTGCTGTAGCAAGGCTCTCAGGACCGATATACGACAAATGATGGT	495	
QY	503	TATATCTTCAATGCAATCGATATCATTTGAGACACGGGTGATGCTTATATCGTTCCTCAC	562	
Db	496	TATGTCCTTTAATCCAGCTGATTAATATCGAAGATACGGGTAATGCTTATATGCTTCCTCAT	555	
QY	563	GGCGACCATTAACCATTAATCTCCTAAGAATCAGTTATTCAGCTTAGCGAGTTAGCTCTGCA	622	
Db	556	GGAGGTCACTATCACTAATTTCCCAAAAGCGATTTAATCTGTAGTGAATTAGCAGCAGCT	615	
QY	623	GAAGGCTAATTTGGAATGGGAAGCAGGATCTCGTCTCTTCAAGTTCTAGTTATTAATGCA	682	
Db	616	AAAGCACATCTGGCTGGAAAAAATAATGCAACCGAGTCAGTTAAGCTATTTCTTCAACACCT	675	
QY	683	AATCCAGCTCAACCAA	698	
Db	676	TCTCCATCTCTTCCAA	691	

RESULT 35
US-09-452-599-34/c
; Sequence 34, Application US/09452599
; Patent No. US20020055101A1
; GENERAL INFORMATION:
; APPLICANT: Bergeron, Michel G.
; APPLICANT: Ouellette, Marc
; APPLICANT: Roy, Paul H.
; TITLE OF INVENTION: Specific and U
; TITLE OF INVENTION: To Rapidly det

TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical
TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
FILE REFERENCE: 12287.31
CURRENT APPLICATION NUMBER: US/09/452,599
CURRENT FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: 08/526,840
PRIOR FILING DATE: 1995-09-11
PRIOR APPLICATION NUMBER: 08/304,732
PRIOR FILING DATE: 1994-09-12
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 841
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
JS-09-452-599-34

Query Match 11.9%; Score 272.4; DB 9; Length 841;
Best Local Similarity 92.9%; Pred. No. 1.3e-65;
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;
1956 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCCAGATAATG 2015
841 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCCAGATAATG 783
2016 GTTTTGGTAACGCTAGCGACCATGTTTCAAAGAAACAAAATGGTCAAGTGTACCAATC 2075
782 GTTTTGGTAACGCTAGCGACCATGTTTCAAAGAAACAAAATGGTCAAGTGTACCAATC 723
2076 AAACGGAAAAACCAACGAGGAGAACTCAGACAGAAACCTCAGAGAAACCTCGAGAAACCCCTC 2135
722 AAACGGAAAAACCAACGAGGAGAACTCAGACAGAAACCTCAGAGAAACCTCGAGAAACCCCTC 663
2136 GAGAAGAGAAACCGCAAGGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAAACCCCTC 2187
662 GAGAAGAGAAACCGCAAGGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAAACCCCTC 603
2188 -----ACCAGAGAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 2240
602 AAGATCACCAGAGAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 543
2241 TTGAAGAAAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 2290
542 T---AAGAAAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 496

RESULT 36
JS-10-121-120-34/c
; Sequence 34, Application US/10121120
; Publication No. US20030180733A1
; GENERAL INFORMATION:
; APPLICANT: Bergeron, Michel G.
; APPLICANT: Ouellette, Marc
; APPLICANT: Roy, Paul H.
; TITLE OF INVENTION: Specific and Universal Probes and Amplification
; TITLE OF INVENTION: Primers
; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial
; TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical
; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
; FILE REFERENCE: 12287.31
; CURRENT APPLICATION NUMBER: US/10/121,120
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/452,599
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: 08/304,732
; PRIOR FILING DATE: 1994-09-12
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
JS-10-121-120-34

Query Match 11.9%; Score 272.4; DB 15; Length 841;
Best Local Similarity 92.9%; Pred. No. 1.3e-65;
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;
1956 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCCAGATAATG 2015
841 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCCAGATAATG 783
2016 GTTTTGGTAACGCTAGCGACCATGTTTCAAAGAAACAAAATGGTCAAGTGTACCAATC 2075
782 GTTTTGGTAACGCTAGCGACCATGTTTCAAAGAAACAAAATGGTCAAGTGTACCAATC 723
2076 AAACGGAAAAACCAACGAGGAGAACTCAGACAGAAACCTCAGAGAAACCTCGAGAAACCCCTC 2135
722 AAACGGAAAAACCAACGAGGAGAACTCAGACAGAAACCTCAGAGAAACCTCGAGAAACCCCTC 663
2136 GAGAAGAGAAACCGCAAGGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAAACCCCTC 2187
662 GAGAAGAGAAACCGCAAGGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAAACCCCTC 603
2188 -----ACCAGAGAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 2240
602 AAGATCACCAGAGAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 543
2241 TTGAAGAAAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 2290
542 T---AAGAAAACTCAGAGAGGCTGAAAGTTTACTTTGGAAAAATCCAGAT 496

RESULT 37
US-09-769-736-23
; Sequence 23, Application US/09769736
; Publication No. US20030138775A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089WO
; CURRENT APPLICATION NUMBER: US/09/769,736
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-09-769-736-23

Query Match 9.8%; Score 225; DB 10; Length 1146;
Best Local Similarity 63.5%; Pred. No. 3.7e-52;
Matches 376; Conservative 0; Mismatches 183; Indels 33; Gaps 1;
100 AAATTTGACACCATGATGAAGTCAAGTAAAGGAGGGGATCAACGGCGAAACAATNGTAT 159
162 AAACAAACAGTGGATCAATCAAGTGTCTGCTGAACAGATCGTAGT 221
160 CAAGATTACGATCAAGGTTTGTGACCTCTCATCGAGACCATTCATTACTATATG 219
222 CAATAATTACTACCAAGGTTTGTGACCTCTCATCGAGGTTGTGATGCGATCCTTAATACCA 281
220 CAAGTTCCTTATGATGCCATCATCAGTGAAGAGCTCCCTCATGAAAGATCCGAATATCA 279
282 GAAAGTTCCTTATGATGCCATTCATTGATGAGAGTTGTGATGCGATCCTTAATACCA 341
280 GTTGAAGGATTGAGACATTGTCAATGAAATCAAGGGTGGTTATGTCAATTAAGTAAACGG 339

Db 342 TTTTAAACATCAGACGTTATCAATGAATCTTAGACGGTTACGTTATTAAAGTCAATGG 401
QY 340 TAAATACATCTGTTACCTTAAAGATCAGAGCTCATGCGGTAATATTCGACAAAAGAGA 399
Db 402 CAACTATTATTGTTTACCTCAAGCAGGTAGTAGCGCAAAAACATTCGAACCAACCAACA 461
QY 400 GANTAAACGTCAAGACGAGGAGCGAGTCAATATCAATCAATCAAA----- 443
Db 462 AATTGCTGAGCAAGTAGCCAAAGGAACTAAAGAGAGCTAAAGAAAGGTTTATGCTCAAGT 521
QY 444 -----GAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGGAGGTTA 486
Db 522 GGCCCATCTCAGTAAGAGAGAGTTGGCGGAGTCAATGAGCAAAAAGACAGAGCGCTA 581
QY 487 TACAACGGATGATGGGTATATCTTCAATGATCTGATATCAATCAATGAGCACGCGGTGATGC 546
Db 582 TACTACAGAGGCTGCTATATTTTATGTCGACAGATATCATTTGATGATTTAGGAGATGC 641
QY 547 TTATATCGTTCCTCACGGGACCAATACCATTTCCCTTAAGATGATGTTATCAGCTAG 606
Db 642 TTATTTAGTACCTCATGTTAATCACTATCATTTATTTCTTAAAAAGATTTGTCTCCAAG 701
QY 607 CGAGTTAGTCTGCAGAGGCTATTGGAATGGGAGCAGGATCTGCTCCT 658
Db 702 TGAGCTAGTCTGCACAGCCTACTGGAGTCAAAAACAGGTCGAGGTGCT 753

RESULT 38
US-09-769-736-17
; Sequence 17, Application US/09769736
; Publication No. US20030138775A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089wo
; CURRENT APPLICATION NUMBER: US/09/769,736
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-09-769-736-17

Query Match 9.8%; Score 225; DB 10; Length 2469;
Best Local Similarity 63.5%; Pred. No. 6.1e-52;
Matches 376; Conservative 0; Mismatches 183; Indels 33; Gaps 1;

QY 100 AAACCTTGACACAGATAGTCAAGTAAAGGAGGAGGATCAACGCCGCAACAAATGTTNAT 159
Db 162 AAACAAAAGAGGATCAATCAGTCTGAAAGAGGCACTCTGCTGAAACAGATCTAGT 221
QY 160 CAAGATTACGGATCAAGGTTATGTGACCTCTCATGAGACCAATATCATTTACTAATGG 219
Db 222 CAAAATTACTGACCAAGGTTATGTTTACCTCACACGGTGACCAATATCAATTTTACCAATGG 281
QY 220 CAAGTTTCCTTATGATGCCATCATCATGATGAGAGCTCTCATGAAAGATCCGATATCA 279
Db 282 GAAAGTTCTTATGATGCGATATTTATGTAAGAGTTGTTGATGACCGATCCTAATACCA 341
QY 280 GTTGAAGGATTCAGACATTTGCTCAATGAATCAAGGTTGTTATGTTCAATTAAGGTAAACGG 339
Db 342 TTTTAAACATCAGAGGTTATCAATGAATCTTAGACGGTTATGTTATTAAAGTCAATGG 401
QY 340 TAAATACATCTGTTTACCTTAAAGATCAGAGCTCATGCGGTAATATTCGACAAAAGAGA 399

Db 402 CAACTATTATTGTTTACCTCAAGCAGGTAGTAGCGCAAAAACATTCGAACCAACCAACA 461
QY 400 GATTAAACGTCAAGACGAGGAGCGAGTCAATATCAATCAATCAAA----- 443
Db 462 AATTGCTGAGCAAGTAGCCAAAGGAACTAAAGAGGCTAAAGAAAGGTTTATGCTCAAGT 521
QY 444 -----GAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGGAGGTTA 486
Db 522 GGCCCATCTCAGTAAGAGAGAGTTGGCGGAGTCAATGAGCAAAAAGACAGAGCGCTA 581
QY 487 TACAACGGATGATGGGTATATCTTCAATGATCTGATATCAATCAATGAGCACGCGGTGATGC 546
Db 582 TACTACAGAGGCTGCTATATTTTATGTCGACAGATATCATTTGATGATTTAGGAGATGC 641
QY 547 TTATATCGTTCCTCACGGGACCAATACCATTTCCCTTAAGATGATGTTATCAGCTAG 606
Db 642 TTATTTAGTACCTCATGTTAATCACTATCATTTATTTCTTAAAAAGATTTGTCTCCAAG 701
QY 607 CGAGTTAGTCTGCAGAGGCTATTGGAATGGGAGCAGGATCTGCTCCT 658
Db 702 TGAGCTAGTCTGCACAGCCTACTGGAGTCAAAAACAGGTCGAGGTGCT 753

RESULT 39
US-09-252-088-13
; Sequence 13, Application US/09252088
; Publication No. US20030031682A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RLOUX, Clment
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Jose
; APPLICANT: MARTIN, Denis
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 8331-9002
; CURRENT APPLICATION NUMBER: US/09/252,088
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US/60/075,425
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 5215
; TYPE: DNA
; ORGANISM: group B streptococcus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(122)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(2511)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (367)..(2511)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Complement ((2716)..(2946))
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Complement ((2995)..(3252))
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Complement ((3299)..(3676))
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Complement ((3837)..(4124))
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Complement ((4351)..(5214))
US-09-252-088-13


```
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089wo
; CURRENT APPLICATION NUMBER: US/09/769,736
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-09-769-736-71

Query Match          5.1%; Score 117; DB 10; Length 1455;
Best Local Similarity 54.8%; Pred. No. 1.1e-21;
Matches 334; Conservative 0; Mismatches 245; Indels 30; Gaps 4;

QY 1470 AACGTTTAGGAAACCAATGCGCAATTAACCTACATGATGATGAGATTCAAGTAGCCA 1529
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
257 AACGTAAGGCAACCAATAGCCAGATTGTCTCAGAGTCTCAAGAAATTGAAGAGGCAA 316
QY 1530 AGTTGGGAGGCAAGTACACACAGAGAGCGGTATATCTTTCATCTCTGATATAACCA 1589
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 AAAAAGCTGGTAAATACACAACTCTGATGGTTTACATTTTGTGCTAAGATATTAAAA 376
QY 1590 GTGATGAGGGGATGCTTATGTAATCCACATATGACCCATAGCCACTGGATTAAAAAG 1649
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
377 AAGATACAGGTACAGGTATGTCTTCCATATGACATGATGATGCTGGTACCAAGA 436
QY 1650 ATAGTTTCTGAACTGAGAGAGCGGAGCCAGGCTTATGCTAAAGAGAAGGTTTGA 1709
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
437 AAGATTTATCAGAGTCGGAATTTAAAGCAGCTCAAGAAATTTCTTTCAGGAAAA ----- 489
QY 1710 CCCCTCTTCGACACACCTCAGGATTCAGGAAATCTAGGCAAAAGGAGCAGAGCTA 1769
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
490 -----TCTGAAGCAATCAAGACAAACCAAAAC-----AGTAAACACAGCTCAAGAAA 538
QY 1770 TCTACACCGGTGAAGCAGCTAAGAGGTGCCACTTGTATCGTATGCTTACAACTTTC 1829
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
539 TCTATGAGCAATTGAACCAAGCAATTTGTTAACTGAGATTTATTTTGGAAATTG 598
QY 1830 AATATACTGTAGAAGTCAAAAACGGTAGTTTAACTACCTCATTATGACCAATTACATA 1889
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
599 CACAAGCGACAGACTATAAGATGGTACATTTGTAATTCCTCATAAAGATCATTACCA 658
QY 1890 ACATCAAAATTTGAGTGGTTTGAGAAAGGCGCTTTA-----TGAGGACCTTAAGGGGT 1940
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
659 ATGTGCAATTAATAATGGTTTGTATGAAGAAAAAGATCTTTTAGCTGATTCAGATAAG 718
QY 1941 ATACTCTTGAGGATCTTTTGGGACCTGCTCAGTACTATGTCGAACATCCAAACGAGTC 2000
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 ATTCTTTAAGAGACTATTATGCTACCGCTAAATATTACATGATGACCCAGAAAAAGCTC 778
QY 2001 CGCATTCAGATATGTTTGGTAAAGCTAGCGACCATGTTTCAAGAAACAAAAATGGTC 2060
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
779 CTAA---AGTTGAAGATGGGTAAAGATGCTGAATTTTATAGGAAGAGGACTCTAATA 835
QY 2061 AAGCTGATA 2069
DB ||||| |||||
836 AAGCAGATA 844

RESULT 42
US-10-087-192-463/c
; Sequence 463, Application US/10087192
; Publication No. US20020182586A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 463
; LENGTH: 31124
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(31124)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-463

Query Match          2.5%; Score 58.2; DB 13; Length 31124;
Best Local Similarity 52.2%; Pred. No. 0.00032;
Matches 129; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 2043 AAAGAAACAAAAATGGTCAAGCTGTATACCAATCAACGAAAAAACCAAGCCGAGGAAAC 2102
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18657 ACAAGACCAAAAGAAAAAGACAAAGACAAAGAAAGAAAGAAAGAAAGAAAGAAAG 18598
QY 2103 CTCAGACAGAAAAACCTTGAGGAAGAAAAACCCCTCCAGAAAGAAAAACCGCAAGCGGAAAC 2162
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18597 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 18538
QY 2163 CAGAGTCTCCAAACCAACAGAGAGAACCCAGAGGATCACCAGAGGAATCAGAAGAACCTC 2222
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18537 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 18478
QY 2223 AGTCCGAGACTGAAAAAGGTTGAAGAAAAACCTGAGAGAGGCTGAAGATTCTTGAAGAAA 2282
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18477 GAGAGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 18418
QY 2283 TCCAGGA 2289
DB |||||
18417 AAGAAGA 18411

RESULT 43
US-09-957-956-6
; Sequence 6, Application US/09957956
; Publication No. US20030130215A1
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; TITLE OF INVENTION: Genomic Polynucleotide Fragments From Chromosome 7
; FILE REFERENCE: JR-14, 000-US
; CURRENT APPLICATION NUMBER: US/09/957,956
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/234,422
; PRIOR FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 45980
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-957-956-6

Query Match          2.4%; Score 54.8; DB 10; Length 45980;
Best Local Similarity 52.7%; Pred. No. 0.0038;
Matches 119; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```


SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 269927
LENGTH: 635
TYPE: DNA
ORGANISM: Human
US-10-027-632-269927

Query Match 2.3%; Score 52.6; DB 16; Length 635;
Best Local Similarity 49.1%; Pred. No. 0.00091;
Matches 136; Conservative 1; Mismatches 140; Indels 0; Gaps 0;

2013 ATGGTTTGGTACGCTAGGACCATGTTCAAGAAACAAATGGTCAAGCTGATACCA 2072
506 ATTCTCTTATTGATTCACCTGACTAGCTTAAAGAGAGAGAGAGATGAGAGG 447
2073 ATCAACGGAACCAAGCGGAGGAGAACTCAGACAGAGAAAACCTGAGGAGAAAACCC 2132
446 AAGAAGAGAGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
2133 CTGAGAGAGAGAAACCGCAAGCGGAGAGAAACCGAGTCTCCAAAACCAACAGAGAACAG 2192
386 GAGAAG 327
2193 AGAATCACCAGAGGATCAGAGAACCTCAGTCTGAGACTGAAAGGTTGAAGAAAAC 2252
326 AAGAAGAAAG 267
2253 TGAGAGAGGCTGAAGATTACTTGGAAAAATCCAGGA 2289
266 AAGAAGAAAG 230

RESULT 47

US-10-052-482-70
Sequence 70, Application US/10052482
Publication No. US20040072264A1
GENERAL INFORMATION:
APPLICANT: Engelhard, Eric
APPLICANT: Morris, David
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71087/RMS/DCF
CURRENT APPLICATION NUMBER: US/10/052,482
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 241
SOFTWARE: PatentIn version 3.1
SEQ ID NO 70
LENGTH: 96596
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (19512)..(20471)
OTHER INFORMATION: "n" at positions 19512 to 20471 can be any base
FEATURE:
NAME/KEY: misc feature
LOCATION: (41369)..(41388)
OTHER INFORMATION: "n" at positions 41369 to 41388 can be any base
FEATURE:
NAME/KEY: misc feature
LOCATION: (85221)..(85596)
OTHER INFORMATION: "n" at positions 85221 to 85596 can be any base
US-10-052-482-70

Query Match 2.3%; Score 52.2; DB 12; Length 96596;
Best Local Similarity 50.6%; Pred. No. 0.034;
Matches 126; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

2041 TCAAGAAACAAAATGGTCAAGTGTATACCAATCAACGGAAGAAACCAAGCGAGGAGAA 2100

1723 TCACCTAGGCAACATAGCGAAACCCCATCTGAAAAAGAAAGAAAGAAAGAAAGAA 1782
2101 ACCTCAGACAGAAAAAACCCTGAGGAGAGAAACCCCTCGAGAGAGAAACCCCAAGCGAGAA 2160
1783 AGAAGAAG 1842
2161 ACCAGAGTCTCCAAAACCAACAGAGAGAACCAAGAAATCACCAGAGGAATCAGAAAGAAC 2220
1843 AGAAGAAG 1902
2221 TCAGGTCGAGACTGAAAAGGTTGAAAGAAAACCTGAGAGAGGCTGAAAGATTACTTGGAAA 2280
1903 AGAAGAAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1962
2281 AATCCAGGA 2289
1963 GGAGGAGGA 1971

RESULT 48

US-09-864-761-19262/c
Sequence 19262, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aomic-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 19262
LENGTH: 305

Tue May 4 13:11:55 2004

;; PRIOR APPLICATION NUMBER: DE 10013847.0
;; PRIOR FILING DATE: 2000-03-15
;; PRIOR APPLICATION NUMBER: DE 10019058.8
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: DE 10019173.8
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 540
;; SEQ ID NO 492
;; LENGTH: 5928
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-492

	Query Match	2.2%;	Score 50.6;	DB 13;	Length 5928;
	Best Local Similarity	53.0%;	Pred. No. 0.015;		
	Matches 131;	Conservative	0;	Mismatches 114;	Indels 2; Gaps 1;
QY	2042	CAAAGAAACAAAATGCTAAGCTGATACCAATCAACGGAAACCAAGCGAGGAGAA	2101		
DB	2248	CTAAAAAATAAAAAATCTAAAAAATAAAAAATCTAAAAAACAACCAACCTAAAAAATAAAAA	2189		
QY	2102	CCTCAGACAGAAAACCTGAGGAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAA	2161		
DB	2188	CCTAAAAACAAAAACCTAAAAACAAAAACCTAAAAAATAAAAACTAAAAAATAAAC	2129		
QY	2162	CCAGAGTCTCCAAAACCAACAGAGG--AAGCAGAAGAAATCACAGAGGAATCAGAGAAC	2219		
DB	2128	CTAAAAAATAAAAACTTAAAAACGAAACCTTAAAAATAATAATTAACCTAAAAATAAAC	2069		
QY	2220	CTCAGGTCGAGACTGAAAAGTTGAAGAAAACCTGAGAGAGGCTGAGAGATTACTTGAA	2279		
DB	2068	TAAAAAATAAAACCTTAAAAAATAAAACCTTAAAAATAAAACCTTAAAAATAAAACCTA	2009		
QY	2280	AAATCCA	2286		
DB	2008	AAAAACA	2002		

Search completed: May 1, 2004, 06:37:07
Job time : 954 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

run on: May 1, 2004, 04:16:59 ; Search time 134 Seconds
(without alignments)
3159.907 Million cell updates/sec

title: US-09-765-272A-66

effect score: 4019

sequence: 1 CSVELGRHQGVKKESNRV.....TEKVEKIREARDLLGKIQD 763

coring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

searched: 682709 seqs, 277475446 residues

total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

command line parameters:

MODEL=frame+ p2n.model -DEV=xlh

Q=/cgn2_1/USPTO.spool/US09765272/runat 30042004_081051_11606/app_query.fasta_1.903

DB=Issued Patents NA -QWMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0

LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -WATRIX=blosum62 -TRANS=human40.cdi

LIST=150 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50

MODE=LOCAL -OUTFMT=prc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

USBR=US09765272_@CGN 1 1 56 @runat 30042004_081051_11606 -NCPU=6 -ICPU=3

NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:**

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:**

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:**

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:**

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:**

5: /cgn2_6/ptodata/2/ina/PCUS_COMB.seq:**

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	4013	99.9	2290	3	US-08-961-083-65	Sequence 65, Appl
2	4013	99.9	2290	4	US-09-536-784-65	Sequence 11, Appl
3	3919	97.5	2531	4	US-09-468-656A-11	Sequence 5, Appl
4	3644.5	90.7	2531	4	US-09-468-656A-5	Sequence 94, Appl
5	3641.5	90.6	8195	4	US-08-961-527-94	Sequence 9, Appl
6	2658.5	66.1	2451	4	US-09-468-656A-9	Sequence 55, Appl
7	2649.5	65.9	2389	3	US-08-961-083-55	Sequence 55, Appl
8	2649.5	65.9	2389	4	US-09-536-784-55	Sequence 243, App
9	1377	34.3	2359	4	US-08-961-527-243	Sequence 355, App
10	1202.5	29.9	6867	4	US-08-961-527-192	Sequence 192, App
11	1202.5	29.9	1684	4	US-08-961-527-258	Sequence 258, App
12	1200	29.9				

13	1167.5	29.0	1455	4	US-09-468-656A-7	Sequence 7, Appl
14	1141.5	28.4	1342	3	US-08-961-083-181	Sequence 181, App
15	1141.5	28.4	1342	4	US-09-536-784-181	Sequence 181, App
c 16	512	12.7	841	2	US-08-743-637B-34	Sequence 34, Appl
c 17	512	12.7	841	3	US-08-526-840B-34	Sequence 34, Appl
18	177	4.4	3095	6	5231168-1	Patent No. 5231168
19	165.5	4.1	2232	1	US-08-257-073-4	Sequence 4, Appl
20	165	4.1	2392	4	US-09-216-393B-331	Sequence 331, App
c 21	165	4.1	2393	4	US-09-216-393B-329	Sequence 329, App
22	163	4.1	4200	1	US-08-242-932-1	Sequence 1, Appl
23	163	4.1	4200	1	US-08-714-481-1	Sequence 1, Appl
24	163	4.1	4200	3	US-08-923-992A-1	Sequence 1, Appl
25	163	4.1	4200	5	PCT-US95-06111-1	Sequence 1, Appl
26	160.5	4.0	4522	4	US-08-956-171B-39	Sequence 39, Appl
c 27	159.5	4.0	80161	3	US-09-036-987A-1	Sequence 1, Appl
c 28	159.5	4.0	80161	3	US-09-370-700-1	Sequence 1, Appl
c 29	159.5	4.0	80161	4	US-09-603-207-1	Sequence 1, Appl
30	158	3.9	1278	4	US-09-286-981B-29	Sequence 29, Appl
31	157.5	3.9	1272	4	US-09-286-981B-31	Sequence 31, Appl
32	156.5	3.9	1276	4	US-09-286-981B-30	Sequence 30, Appl
33	156.5	3.9	12665	4	US-08-961-527-134	Sequence 134, App
c 34	154.5	3.8	28473	4	US-08-961-527-83	Sequence 83, Appl
c 35	153	3.8	580073	4	US-08-545-528D-1	Sequence 1, Appl
36	151	3.8	3384	3	US-08-923-992A-5	Sequence 5, Appl
37	151	3.8	10711	4	US-08-961-527-145	Sequence 145, App
38	150.5	3.7	1338	4	US-09-286-981B-26	Sequence 26, Appl
39	150.5	3.7	5468	4	US-09-535-008-66	Sequence 66, Appl
40	150.5	3.7	5471	4	US-09-535-008-1	Sequence 1, Appl
41	150.5	3.7	5477	4	US-09-535-008-74	Sequence 74, Appl
42	150.5	3.7	5480	4	US-09-535-008-70	Sequence 70, Appl
43	150.5	3.7	5564	4	US-09-535-008-68	Sequence 68, Appl
44	150.5	3.7	5567	4	US-09-535-008-64	Sequence 64, Appl
45	150.5	3.7	5573	4	US-09-535-008-76	Sequence 76, Appl
46	150.5	3.7	5576	4	US-09-535-008-72	Sequence 72, Appl
47	150	3.7	3279	3	US-08-446-137B-1	Sequence 1, Appl
48	149.5	3.7	5471	4	US-09-535-008-62	Sequence 62, Appl
49	149.5	3.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl
50	149.5	3.7	4411529	3	US-09-103-840A-1	Sequence 1, Appl
51	149	3.7	3294	3	US-08-923-992A-7	Sequence 7, Appl
52	149	3.7	3492	3	US-08-923-992A-9	Sequence 9, Appl
53	147	3.7	5295	1	US-09-221-017B-27	Sequence 27, Appl
54	146.5	3.6	6605	1	US-08-769-309A-4	Sequence 4, Appl
55	146.5	3.6	6605	3	US-08-994-570-4	Sequence 4, Appl
56	146.5	3.6	6608	4	US-09-220-132-58	Sequence 58, Appl
57	145	3.6	2370	4	US-09-134-000C-1534	Sequence 1534, Ap
58	145	3.6	3711	4	US-09-566-921-126	Sequence 126, App
59	144	3.6	3312	3	US-08-923-992A-3	Sequence 3, Appl
60	143.5	3.6	1338	4	US-09-286-981B-23	Sequence 23, Appl
61	143	3.6	5710	1	US-08-425-061-6	Sequence 6, Appl
62	143	3.6	5710	1	US-08-825-886-6	Sequence 6, Appl
63	143	3.6	5710	4	US-08-989-890-6	Sequence 6, Appl
64	143	3.6	269223	4	US-09-596-002-41	Sequence 41, Appl
65	142.5	3.5	3534	4	US-09-134-001C-2269	Sequence 2269, Ap
66	142	3.5	31096	4	US-08-956-371E-59	Sequence 59, Appl
67	141.5	3.5	3127	4	US-09-620-312D-613	Sequence 61, App
68	141	3.5	9408	4	US-09-418-710-14	Sequence 14, Appl
69	140.5	3.5	3546	4	US-09-540-236-1884	Sequence 1884, Ap
70	140.5	3.5	6744	1	US-08-119-125A-2	Sequence 2, Appl
71	140	3.5	3752	4	US-08-961-527-208	Sequence 208, App
72	140	3.5	30549	4	US-09-134-001C-322	Sequence 322, App
73	139	3.5	4118	1	US-08-119-125A-1	Sequence 1, Appl
74	139	3.5	5008	3	US-09-110-517-1	Sequence 1, Appl
75	139	3.5	14770	4	US-09-220-132-30	Sequence 30, Appl
76	138.5	3.4	3489	2	US-08-728-323A-1	Sequence 1, Appl
77	138.5	3.4	3489	4	US-09-410-399-1	Sequence 1, Appl
78	138.5	3.4	3489	4	US-09-410-399-1	Sequence 1, Appl
c 79	138.5	3.4	32207	2	US-08-770-379-20	Sequence 20, Appl
c 80	138.5	3.4	32207	3	US-08-757-669A-20	Sequence 20, Appl
81	138.5	3.4	32207	3	US-09-230-371A-20	Sequence 20, Appl
82	138	3.4	11907	3	US-08-061-376-4	Sequence 4, Appl
83	138	3.4	14255	1	US-08-320-559-1	Sequence 1, Appl
84	138	3.4	14255	1	US-08-327-392-1	Sequence 1, Appl
85	138	3.4	14255	1	US-08-306-691B-55	Sequence 55, Appl

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text

Alignment Scores:	
Pred. No.:	0
Score:	4013.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	99.95%
DR:	4
Length:	0
Matches:	763
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0
	2390

US-09-765-272A-66 (1-763) x US-09-536-784-65 (1-2290)

QY	1	CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal	20
DB	2	TGTTCTCTATGAACCTGGTCTGTCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTT	61
QY	21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
DB	62	TCATTATATAGATGCTGATCAGCGTGGTCAAAGGCGAAAACTTCACACCATGAAAGTC	121
QY	41	SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr	60
DB	122	AGTAAGAGGGGGGGATCAACGCCGAAACAATNGNNATCAAGATTACGATCAAGGTTAT	181
QY	61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
DB	182	GTGACCTCTCATGGAGACCAATTATCATTAACAATAATGCAAGTTCTTATGATGCCATC	241
QY	81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal	100
DB	242	ATCAGTGAAGAGCTCCTCATCAAAAGATCCGATTAATCATCAGTTGAAGGATTTCAGACATTGTC	301
QY	101	AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrIleLys	120
DB	302	AATGAATATCAAGGGTGGTTATGTCTATTAAAGGTAAACGGTAATAATCATGTGTTACCTTAAG	361
QY	121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
DB	362	GATCGAGCTCATGGGATTAATATTCGGCAAAAGAAGAGATTAAAGCTCAGAAGCAAGNA	421
QY	141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAalaArgAlaGlnGly	160
DB	422	CGCAGTCATATCAATACTCAAGAGCAGATATGCTGTGTGCTGCACCCAGAGCCCAAGNA	481
QY	161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
DB	482	CGTTATATCAACCGGATGATGGGTATATCTTCATGCACTCTGATCATTCATTGAGCACACGGT	541

541 MetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluAlaAla 560
 1622 ATGACCCATAGCATCTGGATTAAAGAGTAGTTTCTCGAGCTGAGAGCGGAGCC 1681
 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
 1682 CAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTGAGGA 1741
 581 AenThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaLysLysVal 600
 1742 AATACTAGAGCGAAAGGAGCAGAGCTATCTACACCGCGTGAACGAGCTAAGAAGGTG 1801
 601 ProLeuAspArgMetProTyrAsnLeuGlnThrValGluValLysAsnGlySerLeu 620
 1802 CCACITGATCGTATGCTTACAACTTCAATATCTGTAGAGTCAAAAACGGTAGTTTA 1861
 621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluThrPheAspGluGlyLeu 640
 1862 ATCATCACTCATATGACCATACCATACATCAATTTGAGTGGTTTGACGAAGGCTT 1921
 641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrVal 660
 1922 TATGAGGCACTTAAGGGTATACTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGTC 1981
 661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
 1982 GAACATCCAAACGAACGCTCGGCANTCAGATAATGGTGTTCGTAAACGCTACCGACCATGTT 2041
 681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700
 2042 CAAAGAAACAAATGGTCAAGCTGATACCAATCAACAAAGGAAACCAACGAGGAGAAA 2101
 701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720
 2102 CCTCAGACAGAAAACCTCAGGAAGAAAACCCCTCGAGAGAGAAAACCGCAAGCGAGAAA 2161
 721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
 2162 CCAGAGTCTCCAAACCAACAGAGGAAACCAAGAGAAATCACCAGAGGAATCAGAGAACCT 2221
 741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLys 760
 2222 CAGTTCGAGACTGAAAGGTTGAAGAAAACCTCAGAGAGGCTGAGATTACTTGGHAAA 2281
 761 IleGlnAsp 763
 2282 ATCCAGGAT 2290

RESULT 3

US-09-468-656A-11

; Sequence 11, Application US/09468656A

; Patent No. 6582706

; GENERAL INFORMATION:

; APPLICANT: Johnson, Leslie S.

; APPLICANT: Adamou, John E.

; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus

; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural

; TITLE OF INVENTION: Motifs

; FILE REFERENCE: 469201-444

; CURRENT APPLICATION NUMBER: US/09/468,656A

; CURRENT FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: 60/113,048

; PRIOR FILING DATE: 1998-12-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 2531

; TYPE: DNA

; ORGANISM: Streptococcus pneumoniae

US-09-468-656A-11

Alignment Scores:

Pred. No.: 0

Length:

2531

Score: 3919.00 Matches: 745
 Percent Similarity: 98.03% Conservative: 3
 Best Local Similarity: 97.64% Mismatches: 9
 Query Match: 97.51% Indels: 6
 DB: 4 Gaps: 1

US-09-765-272A-66 (1-763) x US-09-468-656A-11 (1-2531)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
 Db 58 TGTTCCTATGAGCTTGGACGTTTCAAGCTGGTCAAGATAGAAAGAGTCTTAATCGAGTT 117
 QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
 Db 118 GCTTATATAGTGTGATCAGCTGGTCAAAAGGCAGAAAACTTCACACCATGAAGTC 177
 QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
 Db 178 AGTAAGAGGAGGAGGATCAACGCCGAAACAAATTTGTATCAAGATTACGGATCAAGGTTAT 237
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
 Db 238 GTGACCTCTCATGAGACCATTATCACTACTATAATGGCAAGGTTCCCTTATGATGCCATC 297
 QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
 Db 298 ATCAGTGAAGAGCTCTCATGAAAGATCCGAATATCACTAGTTGAAGGATTCAGACATGTC 357
 QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
 Db 358 AATGAATCAAGGGTGGTTATGTCATTAAGGTAAACCGGTAAATACTATGTTTACCTTAAG 417
 QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
 Db 418 GATCGRCGTCTATGCGGATAATATTCGGAACAAAGAGAGATTAAACGTGAGAGCGAGGAA 477
 QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
 Db 478 CGCAGTCTATATCACTCACTCAAGACGACATATGCTGTGTGTCGAGCCAGAGCCCAAGGA 537
 QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
 Db 538 CGTTTATACAACGGATGATGGGTATATCTTCAATGCATCTGATATCATTCAGGACACGGGT 597
 QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
 Db 598 GATGCTTATATCGTTCTCCACGCGGACCATTAACCATTAACATTCCTCAAGAAATGAGTTATCA 657
 QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220
 Db 658 GCTAGCGAGTTAGCTGCTGAGAGCCCTATTGGATGGGAGAGGAGATCTCGTCTCTTCT 717
 QY 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
 Db 718 TCAAGTTCTAGTTATATGATAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTG 777
 QY 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
 Db 778 ACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCCTTTTACGTGAA 837
 QY 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
 Db 838 TTGATGCTAAACCCCTTATCAGAACCCGACGAGAGGTAGCTGTCCCTCATGTTAACCATTACCAC 897
 QY 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
 Db 898 GCGCAATCAAGTCCGAACCGCCAGAGGTAGCTGTCCCTCATGTTAACCATTACCAC 957
 QY 301 PheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320
 Db 958 TTTATCCCTTATGAACAAATGTCTGAATTTGGAAAAACGAATTCGTCTGTTATTTCCCTT 1017
 QY 321 ArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSer 340

```
1018 CGTTATCGTTCAACATTGGGTACCAGATTCCAGACCAGAGAACCAAGTCCACAAACG 1077
|||||
341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
|||||
1078 ACTCCAGAACCTAGTCCAGT-----CCGCAACACAGCTCCAGCAAT 1119
|||||
361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
|||||
1120 CCAATTGATGGGAATTGGTCAAGAAAGCTGTTCCGAAAAGTAGCGCATGTTATGCTTT 1179
|||||
381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
|||||
1180 GAGGAGAATGGAGTTCTCGTTATATCCCGACCAAGGATCTTTCCAGCAAAAACAGACGA 1239
|||||
401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420
|||||
1240 GGCATTGATAGCAACTGCCAAGCAGGAAAGTTTATCTCATAGCTAGGAACTAGAAA 1299
|||||
421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
|||||
1300 ACTGACCTCCCATCTAGTCATGAGAAATTTTACAATAAGGCTTATGACTTACTAGCAAGA 1359
|||||
441 IleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
|||||
1360 ATTCCACCAAGATTACTTGATATTAAGSTCGCAAGTTGATTTGAGGCTTTGGATAAC 1419
|||||
461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
|||||
1420 CTGTTGGAACGACTCAAGATGCTCAAGTGAATAGTCAAGTATGAGTGGAAATATCTT 1479
|||||
481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
|||||
1480 GCCTTCTTAGCTCCGATTCTGTCATCCAGACGTTTAGGAAAAACCAATGCGCAATTAAC 1539
|||||
501 TyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
|||||
1540 TACACTGATGATGAGATTCAAGTAGGCAAGTCTGGCAGGCAAGTACACAGCAGAACGCGT 1599
|||||
521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
|||||
1600 TATATCTTTGATCCTCGTATATAACCAAGTAGTGGGGGATGCCCTATGTAATCTCCACAT 1659
|||||
541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAla 560
|||||
1660 ATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGACGCC 1719
|||||
561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
|||||
1720 CAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGGA 1779
|||||
581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysVal 600
|||||
1780 AATACTGAGCAAAAGAGCAGAGCTATCTTACAAACCGMGTTGAAGCAGCTTAAGAGGTG 1839
|||||
601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
|||||
1840 CCATTGATCGTATGCTTACATCTTCAATATATCTAGAGTCAAAAACGGTAGTTTA 1899
|||||
621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640
|||||
1900 ATCATACCTCATTTATGACCATTAACCAATCAATCAAAATTTGAGTGGTTTACGAAGGCGCT 1959
|||||
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrVal 660
|||||
1960 TATGAGCACCTTAAGGGGTATACTCTTTGAGATCTTTTGGCGACTGTCAAGTACTATGTC 2019
|||||
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
|||||
2020 GAACATCCAAACGACGCTCCGATTCAGATTAATGGTTTGGTAAACGCTAGGACCATGTT 2079
|||||
681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluLys 700
|||||
```

```
Db 2080 CAAAGAAACAAAATGTCAGCTGATACCAATCAACCGAAAAACCAAGCGAGGAGAAA 2139
Qy 701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720
Db 2140 CCTCAGACAGAAAAACCTGAGAAAGAAAACCCCTCGAGAAGAGAAAACCGCAAGCGAGAAA 2199
Qy 721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
Db 2200 CCAGAGTCTCCAAACCAACACAGAGAACCAAGAATCACCAGAGGAATCAGAGAACCT 2259
Qy 741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLys 760
Db 2260 CAGTCCAGACTGAAAAGGTTGAAGAAAACTGAGAGAGGCTGAAGATTTACTTGGAAAA 2319
Qy 761 IleGlnAsp 763
Db 2320 ATCCAGGAT 2328

RESULT 4
US-09-468-656A-5
; Sequence 5, Application US/09468656A
; Patent No. 6582706
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural
; FILE REFERENCE: 469201-444
; CURRENT APPLICATION NUMBER: US/09/468,656A
; CURRENT FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-468-656A-5

Alignment Scores:
Pred. No.: 0 Length: 2531
Score: 3644.50 Matches: 698
Percent Similarity: 92.41% Conservative: 20
Best Local Similarity: 89.83% Mismatches: 40
Query Match: 90.68% Indels: 19
DB: 4 Gaps: 4

US-09-765-272A-66 (1-763) x US-09-468-656A-5 (1-2531)
Qy 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 58 TGTTCTATGAACTTGGTCGTCACCAAGCTGGTCAGTTAGAAAGTCTATTCAGATT 117
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATATAGATGGTATCAGCTGCTCAAAAGGCAGAAAACTTCACACCATCAAGTCAAGTTC 177
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGTAAGAGGAGGGGATCAACGCCCAACAAATCGTCATCAAGATTACGGATCAAGGTTAT 237
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTGACCTCTCATGGAGACCATTTATCATATAATGGCAAGGTCCCTTATGATGCAATC 297
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 ATCAGTGAAGAGCTCTCTCATGAAGATCCGATTTATCAGTTCAAGGATTCAGACATTGTC 357
Qy 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
|||||
```

358 AATGAAATCAAGGGTGGTTATCTTATCAAGGTAGATGGAATACTATCTTACCTTAAG 417
121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
418 GATGAGCTCATCGGATATATTTGGACAAAGAGAGATTAAAGCTCAGAGCAGGAA 477
141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAlaGln 159
478 CACAGCTAATACACGGGGTGGTTCTTAAGCATCAAGCAGTAGTTGTCAGCCAGGCCAA 537
160 GlyArgThrThrAspAspGlyTrillePheAsnAlaSerAspIleIleGluAspThr 179
538 GGACGCTATACAAACGATGATGTTATATCTTCAATGCATCTGATATCAATGAGGACCG 597
180 GlyAspAlaTrilleValProHisGlyAspHisTrillePheTrilleProLysAsnGluLeu 199
598 GGTGATGCTTATCTCTCCACGGCGCAATACCATTAACATTCCTAAGAATGAGTTA 657
200 SerAlaSerGluLeuAlaAlaAlaGluAlaTrillePheAsnGlyLysGlnGlySerArgPro 219
658 TCAGCTAGCGATGAGTCTGTCAGAGGCTTATGGAATGGGAAGCAGGATCTCGTCT 717
220 SerSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn 239
718 TCTTCAAGTCTAGTTATATATGCAATCCAGCTCAACCAAGATTGTACAGAGAACCAAT 777
240 LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259
778 CTGAGTGTCACTCCAACTTATCATCAAAATCAAGGGGAACAAATTCAGGCCCTTTACGT 837
260 GluLeuTrilleValProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279
838 GAATTTGTATGCTAAACCTTATCAGAACGCCATGTGGAATCTGATGGCTTATTTGAC 897
280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299
898 CCAGCGCAATCACAAGTCGAACCGCGCAGAGGTGTAGTGTCTCCCTCATGTGAACCAAT 957
300 HisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIlePro 319
958 CACTTTATCCCTTATGAACAAATGCTGAATTTGGAACAAAGCAATGCTGTATTTATCC 1017
320 LeuArgTyrArgSerAsnHisTrilleValProAspSerArgProGluGlnProSerProGln 339
1018 CTTCGTATTCGTTCAACCATTTGGTACAGATTCAGACAGAACCAACAGTCCACAA 1077
340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359
1078 TCGACTCCGGAACCTAGTCCAAAGTCGCAACCTGCAACCAATCTCAACAGCTCCAAAGC 1137
360 AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal 379
1138 AATCCAATTTGATGAGAAATGGTCAAGAGAGCTGTTTCAAGAGTAGCGGATGTTATGTC 1197
380 PheGluGluAsnGlyValSerArgTrilleProAlaLysAspLeuSerAlaGluThrAla 399
1198 TTGAGGAGAAATGGAGTTCTCTGTTATATCCAGCCAAAGTCTTTTCCAGCAGAAACAGCA 1257
400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLysLeuGlyAlaLys 419
1258 GCAGCAATTTGATAGCAAACTGGCCAGCAGCAGGAAGTTTATCTCATAGCTAGGAGCTAAG 1317
420 LysThrAspLeuProSerSerArgGluPheTyrAsnLysAlaTrillePheAspLeuAla 439
1318 AAAAAGTACCTCCATCTAGTATGATGAGAAATTTTACAATTAAGGCTTATGCTTACTAGCA 1377
440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp 459
1378 AGAATTCACCAAGATTACTTGTATATAAAGTGCACAAAGTTGATTTGAGGCTTTGAT 1437
460 AsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIle 479
1438 AACCTGTTGGAAGCATCAAGGATGTCCCAAGTGATTAAGTCAAGTTAGTGGATGATTT 1497

QY 480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499
Db 1498 CTTGCCCTTCTAGCTCCGATTTCGTATCCAGAAAGTTAGGAAAAACCAATTCGCAAAAT 1557
QY 500 ThrTyrThrAspAspGluIleGlnValAlaLysLysLeuAlaGlyLysTyrThrThrGluAsp 519
Db 1558 ACCTACACTCATGATGAGATTCAAGTAGCCCAAGTTGGCAGGCAAGTACACACAGAGAC 1617
QY 520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTrilleValThrPro 539
Db 1618 GGTATATATCTTGTATCTCTGATATATACCAAGTATGAGGGGATGCCTATGTAATCTCA 1677
QY 540 HisMetThrHisSerHisTrilleLysLysAspSerLeuSerGluAlaGluAlaAla 559
Db 1678 CATATGACCCATAGCCACTGGATTAAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCA 1737
QY 560 AlaGlnAlaTrilleValLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer 579
Db 1738 GCCCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCTTCACAGACCATCAGGATTCA 1797
QY 580 GlyAsnThrGluAlaLysGlyAlaGluAlaTrilleTyrAsnArgValLysAlaAlaLysLys 599
Db 1798 GGAATATCTGAGGCAAAAGGAGCAGAGCTATCTACAAACCGGTGAAGACGCTTAAGAG 1857
QY 600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619
Db 1858 GTGCCACTTGTATGATGATGCTTACAACTTCAATATCTGTAAGTCAAAACCGTACT 1917
QY 620 LeuIleIleProHisTyrAspHisTyrHisAsnLysLysPheGluTrillePheAspGluGly 639
Db 1918 TTAATCATACCTCATTTATGACCAATTCATTAACATCAAAATTTAGTGTGTTTGCAGGAGC 1977
QY 640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyr 659
Db 1978 CTTTATGAGGCACTTAAGGGGTATCTCTTGGAGATCTTTTGGCGACTGTCAAGTACTAT 2037
QY 660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679
Db 2038 GTCGACATCCAAACGAAAGTCCGCAATTCAGATATGTTTGGTAAACGCTAGCGACCAT 2097
QY 680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699
Db 2098 GTTCGTAATAATAAG-----GTAGACCAAGACAGTAAACCTGATGAAGAT 2142
QY 700 LysProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGlu 719
Db 2143 AAGAAACATGATGAAGTAAGTAGGACCAACTCACCCCTGAATCTGTGAAAGAGAGAAATCAC 2202
QY 720 LysProGluSerPro-----LysProThrGluGluProGluGluSer 733
Db 2203 GCTGTTTAATCTTTCAGCAGATAATCTTTATTAACCAAGCACTGATACGGAAGAGACA 2262
QY 734 ProGluGluSerGluGlu-----ProGlnValGluThrGluLys 746
Db 2263 GAGCAAGAAAGCTGAAGATACCACAGATCAGGCTGAAATTCCTCAAGTAGAGAATTCGTGT 2322
QY 747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763
Db 2323 ATTAACGTAAGTAGCAGATGCGGAGGCTTGTCTAGAAAAGTAACAGAT 2373

RESULT 5

US-08-961-527-94
; Sequence 94, Application US/08961527
; Patent No. 6420135

GENERAL INFORMATION:
; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 8195 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-94

Alignment Scores:
Pred. No.: 0 Length: 8195
Score: 3641.50 Matches: 698
Percent Similarity: 92.41% Conservative: 20
Best Local Similarity: 89.83% Mismatches: 40
Query Match: 90.61% Indels: 19
DB: 4 Gaps: 4

US-09-765-272A-66 (1-763) x US-08-961-527-94 (1-8195)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
DB 5659 TGTTCCTATGAATGGTCGTCCACCAAGCTGCTCAGGTTTAAAGAGAGTCTAATCGAGTT 5718
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProaspGluVal 40
DB 5719 KCTTATATAGATGGTGATCAGCTGCTCAAAAGGCGAGAAACTTGCACACAGATGAAGTC 5778
QY 41 SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr 60
DB 5779 AGTAAGAGGAGGGGATCAACCGGACAAATCGTCATCAAGATTACGGATCAAGGTTAT 5838
QY 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
DB 5839 GTGACCTCTCATGGAGACCATTAATCTACTATATAATGGAAGGTCCTTATGATGCCATC 5898
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
DB 5899 ATCAGTGAAGAGCTCCTCATGAAGATCCGAATTTATCAGTTGAAGGATTCAGACATGTC 5958
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
DB 5959 AATGAATCAAGGTGGTTATGTTATCAAGGTAGATGGAAATCTATGTTTACCTTAAG 6018
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
DB 6019 GATGCAGCTCATCGGATAATATTCGACAAAGAGAGATTAAACGTCAAGAGCAGGAA 6078
QY 141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaArgAlaGln 159
DB 6079 CACAGTCATAATCACGGGGTGGTTCTTAACGATCAAGCAGTAGTTGCAGCCAGAGCCCAA 6138
QY 160 GlyArgTyrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr 179

DB 6139 GGACGCTATACACGAGATGATGTTATATCTTCAATGCACTGATATCATTTAGAGCACG 6198
QY 180 GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu 199
DB 6199 GGTGATGCTTATATCGTTCCTCACGCGACCATTAACCATTAATTAAGAAATGAGTTA 6258
QY 200 SerAlaSerGluLeuAlaAlaGluAlaTyrTropAsnGlyLysGlnGlySerArgPro 219
DB 6259 TCAGCTAGCAGTTAGCTGCTCAGAGACCTATTGGAATGGGAAGCAGGATCTCGTCCT 6318
QY 220 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn 239
DB 6319 TCTTCAAGTTCTAGTTTAAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCAAT 6378
QY 240 LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259
DB 6379 CTGACTGTCTCACTCCAACTTATCATCAATCAAGGGGAAACATTTCAAGCCCTTTACGT 6438
QY 260 GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279
DB 6439 GAATGTATGCTAAACCTTATCAGAACGCGATGTGGAATCTGATGGCCTTATTTTCGAC 6498
QY 280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299
DB 6499 CCAGCGCAATCACAGTCGAAACCGCAGAGGTGAGTCTCCCTCATGGTAACCATATC 6558
QY 300 HisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIlePro 319
DB 6559 CACTTATCCCTTATGAAACAATGCTGAATCGGAAAAACGAATGCTCGTATTTATCCC 6618
QY 320 LeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGln 339
DB 6619 CTTTCGTTATCGTTCAAAACCATTTGGTACCAAGATTCAAGACAGAACCAACCAAGTCCCAA 6678
QY 340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359
DB 6679 TCGACTCCGGAACCTAGTCCAGTCCGCAACCTCGCAAAATCTCAACAGCTCCCAAGC 6738
QY 360 AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal 379
DB 6739 AATCCAAATGATGAGAAATTTGGTCAAGAGAGCTGTTGAAAAGTAGGCGATGTTATGTC 6798
QY 380 PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla 399
DB 6799 TTTGAGGAGAAATGGAGTTTCTCGTTATATCCAGCCCAAGATCTTTCAGCAGAACAGCA 6858
QY 400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGluAlaLys 419
DB 6859 GCAGGCAATGATAGCAACCTGGCCCAAGCGAAAGTTTATCTCATAGCTAGGAGCTAAG 6918
QY 420 LysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAla 439
DB 6919 AAACTGACCTCCCACTAGTATCGAGAAATTTTACAAAGGCTTATGACTTACTAGCA 6978
QY 440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp 459
DB 6979 AGAATTCACCAAGATTTTACTTTGATAATAAAGTCCGAAAGTGTATTTTGGAGCTTTGGAT 7038
QY 460 AsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIle 479
DB 7039 AACCTGTGGAAACGACTCAAGGATGTCYCAAGTATTAAGTCAAGTATGATGATATT 7098
QY 480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499
DB 7099 CTTCGCTCTCTTAGCTCCGATTCGCATCCAGAACGTTTGGAAACCAAAATTCGCGAAT 7158
QY 500 ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp 519
DB 7159 ACCTACACTGATGATGAGATTCAAGTACCGCAAGTTGGCAGGCAAGTACACACAGAGAC 7218
QY 520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro 539

286 ArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGlu 305
955 CGAACAGCAGTAGAGGTGTTGCGAGTGCACACGAGATCATTAACCACTTCATCCCTTACTCT 1014
306 GlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerIle 325
1015 CAATGTCGTAATTGGAAGAAGCAATCGCTGATATTATTCCTCCCTTCGTTATCGTTCAAA 1074
326 HisTyrValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSer 345
1075 CATTCGGGTACAGATTCAAGGCCAGAAACAACAAGTCCACACCGACTCCGGAACTAGT 1134
346 ProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLys 365
1135 CCAGCCCGCAACCTGCACCAATCTTAAATATAGACTCAAT-----TCTTCT 1182
366 LeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyVal 385
1183 TTGGTTAGTCAGCTGTTACGAAAGTGGGGAAGGATATGTTATCGAAGAAAGGCGATC 1242
386 SerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys 405
1243 TCTGTTATGTTCTTGGCAAGATTACATCTGAACCTGTTAAATCTTGAAGCAAG 1302
406 LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSer 425
1303 TTATCAAAACAAGAGAGTGTTCACACACTTTAACTGCTTAAATAAGAAATAATGTTCCTCT 1362
426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445
1363 CGTGACCAAGAAATTTATGATAAGCATATATCTGTTAACTGAGGCTCATAAAGCCTTG 1422
446 LeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeu 465
1423 TTTGNAATATAGGTCGTAATCTGATTTCCAGCCTTAGACAATATTATAGACGCTTG 1482
466 LysAspVal**SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaPro 485
1483 AATGATGATGACTAATAAGAAAATTTGATGATGATTATTTGCTTCTCCTAGCACCA 1542
486 IleArgHisProGluArgLeuGlyLysProAlaGlnIleThrTyrThrAspAspGlu 505
1543 ATTACCATTCAGACGCACTTGGCAACCAATCTCAAAATGAGTATACTGAAGACGAA 1602
506 IleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspPro 525
1603 GTTCGTAATGCTCAATAGCTGATAGTATACACGTCAGATGGTTTACATTTTGTATGAA 1662
526 ArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHis 545
1663 CATGATATAATCAGTGATGAAGGAGATGCATATGTAACGCTCATATGGCCATAGTCAC 1722
546 TrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAlaLys 565
1723 TGGATTGAAAAGATAGAGCTTCTGTAAGGAAAAAGTTGCAAGCTCAAGCCCTATATAA 1782
566 GluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLys 585
1783 GAAAAGATATCCTACCTCCATCTCCAGACGAGATGTTAAAGCAAAATCCAACTGGAGAT 1842
586 GlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMet 605
1843 AGTCAGCAGCATTTTACAATCGTGTGAAAGGGAAGAAAGAAATTCACCTCGTTCGACTT 1902
606 ProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyr 625
1903 CCATATATGTTGAGCATACAGTTGAGGTTAAACCGGTAATTTGATTATTCCTCATAG 1962
626 AspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLys 645
1963 GATCATTAACCAATAATTATTAATTTGCTTGGTTTGTATGATCACACATACAAAGCTCAAT 2022

646 GlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665
2023 GGCTATACCTTGAAGATTTGTTGGACGATTAAGTACTACGTAGAACACCCCTGACGAA 2082
666 ArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn 685
2083 CGTCCACATCTTAATGATGGGCAATCCAGTCAGCATGTGTAGGCAAGAAAGAC 2142
686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
2143 CACAGTGAAGATCCAAATAAGAACTTCAAAGCGGATGAAGAG----- 2184
706 ProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725
2185 CCAGTAGAGGAAACA----- 2199
726 ProThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGlu 745
2200 -----CCTGCTGAGCCAGAGTCCCTCAAGTAGAGACTGAA 2235
746 LysValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763
2236 AAAGTAGAAGCCCACTCAAAGAGCAGAGATTTCCTTCGGAAGTAACGGAT 2289

RESULT 7
US-08-961-083-55
; Sequence 55, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-55

Alignment Scores:
Pred. No.: 5,96e-243 Length: 2389
Score: 2649.50 Matches: 519
Percent Similarity: 76.71% Conservative: 77
Best Local Similarity: 66.80% Mismatches: 132
Query Match: 65.92% Indels: 49
DB: 3 Gaps: 8

US-09-765-272a-66 (1-763) x US-08-961-083-55 (1-2389)

1019 TGGGTACAGATTCAAGCCGAGAAACAACCAAGTCCACCAACCGACTCCGGAACCTAGTCCA 1078
QY SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
Db GCGCGCAACCTGCACCAAACTTAA--ATAGACTCAAT-----TCTTCTTG 1126
QY ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
Db GGTAGTCAGCTGGTACGAAAGTGGGGAAGGATGATGATTCGAAGAAAGGCGCATCTCT 1186
QY ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
Db CGTTATGCTTTGGCAAGATTTACCACTGTAACCTGTTAAATACTCTGAAAGCAAGTTA 1246
QY AlalysGlnGluSerLeuSerHisLysLysLeuGlyAlaLysLysThrAspLeuProSerSer 426
Db TCAAAACAAGAGAGTGTTCACACACTTTAACTGCTTAAATAAGAAATGTTGCTCTCTCGT 1306
QY AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeu 446
Db GACCAAGAATTTATGATAAGCATATAATCTGTTAACTGAGCTCATAAAGCCTTGTTT 1366
QY AsphenLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLys 466
Db GNAATAAGGGTCGTAATTTCTGATTTCCAAAGCCTTAGACAAATTTATAGAACGCTTGAAT 1426
QY AspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIle 486
Db GATGATCGCTAATAAGAAATTTGGTAGATGATTTATGTCATTCCTAGCACAAT 1486
QY ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506
Db ACCATCCAGAGCGACTGGCAACCAATTTCTCAATTTGAGTATCTAGAGAGAGTT 1546
QY GlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArg 526
Db CGTATTTGCTCAATTAGCTGATAAGTATACACGTCAGATGGTTACATTTTGTGAACAT 1606
QY AspileThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr 546
Db GATATAATCAGTGATGAAGGAGATCATATGTAACGCTCATATGGCCCTCATGCTACTG 1666
QY IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGlu 566
Db ATTGAAAAGATAGCTTCTGATGAAGAAAGTTGAGCTCAAGCTTATCTACTAAGAA 1726
QY LysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGly 586
Db AAAGTATCTTACCTCCATCTCCAGAGCGAGATGTTAAAGCAATCCAACTGGAGATAGT 1786
QY AlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMetPro 606
Db GCAGCAGCTATTACAACTGTTGAAGGGGAAAAACGAATTCACCTGCTCCACTTCCA 1846
QY TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAsp 626
Db TATATGTTGAGTACAGTTCAGGTTAAACCGGTAATTTGATTTCTCTCAAGGAT 1906
QY HisTyrHisAsnIleLysPheGluThrPheAspGluGlyLeuTyrGluAlaProLysGly 646
Db CATTAACATAATTAATTTGCTGTTGATGATCACACATACAAAGCTCCAAATGGC 1966
QY TyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProLeuGluArg 666
Db TATACCTTGAAGATTTCTTTCGCGAGTAACTAGTACTACGTAGAACACACCTGACGACGT 2026
QY ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGly 686
Db CCACATCTTAATGATGATGGGCAATCCAGTACGATGTTGTTAGGCAAGAAAGACCAC 2086
QY GlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysPro 706
Db AGTGAAGATCCAAATAGAACTTCAAGCGGATGAAGAG-----CCA 2128

QY 707 GluGluGluThrProArgGluGluGluProGlnSerGluLysProGluSerProLysPro 726
Db 2129 GTAGAGAAACA-----2140
QY 727 ThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLys 746
Db 2141 -----CCTGCTGAGCCAGAGCCCTCAAGTAGAGACTGAAAAA 2179
QY 747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763
Db 2180 GTAGAGCCCACTCAAGAGCAGAGAGTTTGTCTTGCGAAGTAACGGAT 2230

RESULT 8
US-09-536-784-55
; Sequence 55, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340B3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:

US-09-536-784-55
Alignment Scores:
Pred. No.: 5.96e-243 Length: 2389
Score: 2649.50 Matches: 519
Percent Similarity: 76.71% Conservative: 77
Best Local Similarity: 66.80% Mismatches: 132
Query Match: 65.92% Indels: 49
DB: 4 Gaps: 8

US-09-765-272A-66 (1-763) x US-09-536-784-55 (1-2389)
QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21
Db 2 TCCTACAGTTGGGACTGTATCAAGCTAGACGGTT---AGGAAATAATCGTGTTC 58
QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSer 41
Db 59 TATATAGTGAACAAAGCAGCGCAAAACCGGAGAAATTTGACTCTCTGATGAGGTTAGC 118

QY 42 LysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyrVal 61
Db 119 AAGCGTGAAGGAATCAATCGTCAAGCAATCGTCATCAAGATAACAGACCAAGGCTATGTC 178
QY 62 ThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIleIle 81
Db 179 ACTTCACATGGCGACCACTATCATTTATTAATGGAAGTTCTTATGACGCTATCATC 238
QY 82 SerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsn 101
Db 239 AGTGAAGAATTACTCATGAAGATCCAACTATAGCTAAAGATGAGGATATGTTAAT 298
QY 102 GluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAsp 121
Db 299 GAGGTCAGGGTGGATATGTTATCAAGGTAGTAGGAAATACTATGTTTACCTTAAGGAT 358
QY 122 AlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnLysArg 141
Db 359 GCTGCCCAACGCGATAACGTCGTACAAAGAGGAAATCAATCGACAAAACAGAGCAT 418
QY 142 SerHisAsnHisAsnSer-----ArgAlaAspAsnAlaValAlaAlaArgAla 158
Db 419 AGTCAACATCGTGAAGTGAATCTCAAGAACGATGGTCTGTTGCTTGGCCTGGCAGTTCG 478
QY 159 GlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAsp 178
Db 479 CAAGGAGCGCTACTACAGATGATGTTATCTTTAATGCTTCTGATATCATAGAGAT 538
QY 179 ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGlu 198
Db 539 ACTGCTGATGCTTATATCGTCTCATGAGGATCATTAACATTAACATTAACATTAACAT 598
QY 199 LeuSerAlaSerGluLeuAlaAlaGluAlaTyrTipAsnGly-----213
Db 599 TTATCAGCTAGCGAGTTGGTCTGTCGAGAGCCCTCTCTATCTGTCGAGAAATCTGTCA 658
QY 214 -----LysGlnGlySerArgProSerSerSerSerSerSerSerSerSerSer 227
Db 659 AATTCAAGAACCTATCGCCGACAAATAGGATAACACTTCAAGAACAACTGGGTACCT 718
QY 228 AsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrHis 247
Db 719 TCTGTGAAGCAATCCAGGAATCAATAATACTAACACAGCAACACACACACACACAGT 778
QY 248 Gln---AsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu 266
Db 779 CAAGCAAGTCAAGTAATGACATTTGATGTTCTTGAACAGCTCTACAAACTGCCTTTG 838
QY 267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286
Db 839 AGTCAACGACATGTAGATCTGATGCGCTTGTCTTGTATCCAGCAACATCAAGTCA 898
QY 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
Db 899 ACAGCTAGAGGTGTCAGTGCACACGAGGATCATTAACATTAACATTAACATTAACAT 958
QY 307 MetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHis 326
Db 959 ATGCTCGAATTGGAAGAACGAATCGCTCGTATTAATTAATTAATTAATTAATTAATTA 1018
QY 327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346
Db 1019 TGGGTACCAAGATTCAAGGCCAGAACCAACAGTCCACACCGACTCCGGAACCTAGTCCA 1078
QY 347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
Db 1079 GSCCGCACTCGACCAATCTTAAAT---ATAGACTCAAAAT-----TCTTCTTG 1126
QY 367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluAsnGlyValSer 386
Db 1127 GTTAGTCTGCTGTCGAAAGTTGGGAGGATATGTATTCGAGAAAGGCGCATCTCT 1186

387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
1187 CTTATGCTCTTCGCGAAAGATTTTACCACCTGCAAACTGTTAAAAATCTTGAAGAAGTTA 1246
407 AlaLysGlnGluSerLysLeuGluValAlaLysLysThrAspLeuProSerSer 426
1247 TCAAAACAGAGAGTGTTCACACACTTTAACTGCTAAAAAGAAAAGTGTCTCCTCGT 1306
427 AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 446
1307 GACCAAGAATTTTATGATAAGCATATACTGTAACTGAGGCTCATAAAGCCCTTGTT 1366
447 AspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLys 466
1367 GNAATAAGGCTGCTAATCTGATTTCCAGCCCTTAGACAAATTTATTAGACGCTGAAT 1426
467 AspVal**SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle 486
1427 GATGAATCGACTAATAAGAAAATTTGGTAGATGATTTATTTGGCATTTCTAGCCCAATT 1486
487 ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506
1487 ACCCATCCAGAGCGACTTGCACAAACCAATTCAAATTCAGTATATCTGAAGACGAGTT 1546
507 GlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArg 526
1547 CGTATTGCTCAATTAGCTGATAAGTATACAAACGTCAGATGTTTACATTTTGTGAAACAT 1606
527 AsnIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr 546
1607 GATATAATCAGTGAAGAGAGATGATATGTAAACGCTCATATGGGCCCATAGTCACTGG 1666
547 IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGlu 566
1667 ATTGGAAGAAGATAGCTTCTGATAGGAAGAAAGTTCAGCTCAAGCTATCTACTAAAGAA 1726
567 LysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGly 586
1727 AAAGGTATCTCATCTCCAGACGCGAGATGTTAAAGCAATCCAACTGGAGATAGT 1786
587 AlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMetPro 606
1787 GCAGCAGACTATTACAACTGCTGTGAAGGGGCAAAACGAATTCACCTCGTTCGACTTCCA 1846
607 TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleProHisTyrAsp 626
1847 TATATGTTGAGCATACAGTTGAGGTTAAACCGTAAATTTGATTTCTCTCATAGGAT 1906
627 HisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAlaProLysGly 646
1907 CATACCAATAATTAATTTGCTGTTGATGATCCACATACAAAGCTCCAAATGGC 1966
647 TyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg 666
1967 TATACCTCGGAAGATTTGTTGCGAGATTAAGTACTACTAGAACACCTCGTAGCAACGT 2026
667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGly 686
2027 CCACATCTTAATGATGGGCAATGCCAGTGAGCATGTTGTTAGCAAGAAGACCAAC 2086
687 GlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysPro 706
2087 AGTGAAGATCCAAATAAGAACTTCAAGCGGATGAAGAG-----CCA 2128
707 GluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLysPro 726
2129 GTAGAGCAACA----- 2140
727 ThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLys 746
2141 -----CCTGCTGACCCAGAGCTCCCTCAAGTAGAGACTGAAAAA 2179
747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763

2180 GTAGAAGCCCACTCAAGAGCAGAAAGTTTTCCTTGCAGAAAGTAACGGAT 2230
RESULT 9
US-08-961-527-243
Sequence 243, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 243:
SEQUENCE CHARACTERISTICS:
LENGTH: 2359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-243
Alignment Scores:
Pred. No.: 1.98e-221 Length: 2359
Score: 2424.00 Matches: 465
Percent Similarity: 97.50% Conservative: 3
Best Local Similarity: 96.88% Mismatches: 6
Query Match: 60.31% Indels: 6
DB: 4 Gaps: 1
US-09-765-272A-66 (1-763) x US-08-961-527-243 (1-2359)
Qy 1 CysSerTyrGluLeuGluArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 937 TGTTCCTATGAGCTTGGACGTTTACCAGCTGTCTCAGGATAGAAAGAGTCTAATCGAGTT 996
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 997 GCTTATATAGTGTGATCAGCTGCTCAAAAGGCAGAAACTTCACACCAGATGAAGTC 1056
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 1057 AGTAAGAGGAGGGGATCAACGCCGAACAAAATGTTTATCAAGATTACGATCAAGGTTAT 1116
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
Db 1117 GTGACCTCTCATGTGAGACCATTTATCTACTATATATGCAAGGTTTCTTTATGATGCCATC 1176
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100

Db 1177 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAAATATCAGTTGAAGGATTACAGACTTGTCTC 1236
Qy 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyTyrTyrValTyrLeuLys 120
Db 1237 AATGAAATCAAGGCTGTTATGTCTATTAAGGTAAACGGTAATAACTATGTTTACCTTAAG 1296
Qy 121 AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 1297 GATGCAGCTCATGCGGATAATATTCGGAACAAAGAGAGATTAAACGTCAGAAAGCAGAA 1356
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
Db 1357 CGAGCTCATATCACTCAAGAGCAGATATGCTGTGTCAGCCAGGACCCAAAGGA 1416
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 1417 CGTTATACACGGATGATGGGTATATCTTCAATGCACTGATATCAITGAGACACGGGT 1476
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 1477 GATGCTTATATCGTTCCTCACGGCGACCAATACCAATTACATTCCTCAAGATGAGTTATCA 1536
Qy 201 AlaSerGluLeuAlaAlaGluAlaTyrTyrPheGlnGlyLysGlnGlySerArgProSer 220
Db 1537 GCTAGCGAGTTAGCTGCTGCAGAGCCCTATGGAAATGGGAAGCAGGAGATCTCGTCCCTCT 1596
Qy 221 SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
Db 1597 TCAAGTTCTAGTTATATGCAATCCAGCTCAACAGATTTGTACAGAAACCAATCTG 1656
Qy 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
Db 1657 ACTGTCACCTCAACTATCATCAAAATCAAGGGGAAACATTTCAAGGCTTTTACGTGAA 1716
Qy 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
Db 1717 TTGTATGCTAAACCTTATCAGAACGCCATGTGAATCTGATGCTGATTTTTCGACCA 1776
Qy 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
Db 1777 GCGCAATCAACAGTCGACCGCAGAGGTAGTGTCCCTCATGTGTAACCATTACAC 1836
Qy 301 PheIleProTyrGluGlnMetSerGluLeuLysArgIleAlaArgIleIleProLeu 320
Db 1837 TTTATCCCTTATGAACAAATGCTGTGAATGGAAAAAGCAATGCTCGTATTATTTCCCTT 1896
Qy 321 ArgTyrArgSerHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340
Db 1897 CGTTATCGTTCAAAACCATGCGTACCAAGATTCAAGACCCAGAAAGAACCAAGTCCCAACCG 1956
Qy 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
Db 1957 ACTCCAGAACTAGTCCAAAGT-----CCGCAACCGACCTCCAAAGCAAT 1998
Qy 361 ProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
Db 1999 CCAATTGATCAGAAATTTGGTCAAGAGAGCTGTCGAAAGTAGGCGATGGTTATGCTCTT 2058
Qy 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
Db 2059 GAGGAGAATGGAGTTCTCGTTATATCCAGCCAGGATCTTTTCAGCAGAAACAGCAGCA 2118
Qy 401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420
Db 2119 GGCATTGATAGCAAACTGGCCCAAGCAGAAAGTTTATCTATAAGTAGTAGGAACTAAGAAA 2178
Qy 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
Db 2179 ACTGACCTCCCATCTAGTATCGAGAAATTTCAATAGGCTTATGACTTACTAGCAAGA 2238
Qy 441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
Db 2239 ATTCAACCAAGATTTACTTGATAATAAAGGTCGACAAAGTTGATTTTGGGCTTTGGATAAC 2298

Qy 461 IeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
Db 2299 CTGTTGGACCACTCAAGGATGCTCAAGTGATAAAGTCAAGTTAGTGAAGATATTCTT 2358
RESULT 10
US-08-961-527-355
; Sequence 355, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunach
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-355
Alignment Scores:
Pred. No.: 4,43e-122 Length: 973
Score: 1377.00 Matches: 266
Percent Similarity: 98.17% Conservative: 2
Best Local Similarity: 97.44% Mismatches: 5
Query Match: 34.26% Indels: 1
DB: 4 Gaps: 0
US-09-765-272A-66 (1-763) x US-08-961-527-355 (1-973)
Qy 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 155 TGTTCCTATAGCTGGAGCTTACCAAGCTGTGTCAGGATGAAGAGAGTCTAATCGATT 214
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 215 GCTTATATAGCTGGTGTGATCAGCTGGTCAAAAGGCGAAGAACTTGACACAGATGAATC 274
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 275 AGTAAGAGGGAGGGGATCAACCGCAACAAATTTGTTATCAAGATTACGATCAAGGTTAT 334
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
Db 335 GTGACCTCTCATGGAGACCATTTATCTACTATAATGGAAGGTTCTCTTATGATGCATC 394

```
81  IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
    |||
    |||
395  ATCAGTGAAGAGCTCCTCATGAAGATCCGAAATTATCATGTTGAAGATTACAGCATGTC 454
    |||
101  AsnGluLysGlyGlyTyrValIleLysValAsnGlyLysTyrValTyrLeuLys 120
    |||
455  AATGAATCAAGGGTGTGTTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 514
    |||
121  AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluLysArgGlnLysGlnGlu 140
    |||
515  GATGAGCTCATCGGATATATATCGACAAAGACAGATTAACCTCAGAACAGGAA 574
    |||
141  ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly 160
    |||
575  CGCAGTCATTAATCATTAATCAAGAGCAGATTAATGCTGTCGAGCCAGAGCCAAAG 634
    |||
161  ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
    |||
635  CGTTATACACGAGTATGAGGATATATCTCAATGCACTCATGATCATGAGGACACGGGT 694
    |||
181  AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
    |||
695  GATGCTTATATCTGCTCTCAGCGGACCATTAATCAATCAATCAATCAATCAATCAAT 754
    |||
201  AlaSerGluLeuAlaAlaAlaGluAlaTyrTriPasnGlyLysGlnGlySerArgProSer 220
    |||
755  GCTAGCGAGTAGCTGCTGCAGAGGCTATGGAATGGGAAGCAGGATCTCGTCTCTCT 814
    |||
221  SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
    |||
815  TCAAGTCTAGTTATATATGCAATATCCAGCTCAACCAAGATTGTGAGAGAACCAATCTG 874
    |||
241  ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
    |||
875  ACTGTCACTCCAACTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 934
    |||
261  LeuTyrAlaLysProLeuSerGluArgHisValGluSer 273
    |||
935  TTGTATGCTAA-CCCTTATCAGAACGCCATGTGGATCT 972
    |||

RESULT 11
US-08-961-527-192/c
; Sequence 192, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
```

```
TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-192

Alignment Scores:
Pred. No.: 4,61e-104 Length: 6867
Score: 1202.50 Matches: 294
Percent Similarity: 54.04% Conservative: 101
Best Local Similarity: 40.22% Mismatches: 187
Query Match: 29.92% Indels: 150
DB: 4 Gaps: 23

US-09-765-272A-66 (1-763) x US-08-961-527-192 (1-6867)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
    |||
    |||
6725 TGTGCTATGCACTTAACACGAGCATGTTGCG---CAGAAAATTAAGCACAATATCGTGTG 6669
    |||
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
    |||
6668 TCTTATGTGATGGCAGCCAGTCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6609
    |||
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
    |||
6608 AGCCAGAAAGAAAGGAATTCAGGCTGAGCAAAATGTAATCAAAATTCAGATCAGGGCTAT 6549
    |||
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
    |||
6548 GTAACGTCACGCGTGACACTATCATTAATTAATGGAAGATTCCTTATGATGCGCTC 6489
    |||
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
    |||
6488 TTTAGTGAAGAACTCTTGATGAAGGATCCAACTATCAACTTAAAGACGCTGATATTGTC 6429
    |||
QY 101 AsnGluLysGlyGlyTyrValLysValAsnGlyLysTyrTyrValTyrLeuLys 120
    |||
6428 AATGAAGTCAAGGGTGTGTTATATCATCAAGTCGATGGAAAAATATTATGTACCTGAAA 6369
    |||
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluLysArgGlnLysGlnGlu 140
    |||
6368 GATGAGCTCATGCTGATTAATGTTTCACTAAGATGAATCAATCTCAAAAACAGAA 6309
    |||
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly 160
    |||
6308 CATGTCAAAAGAT---AATGAGAAGGTTAACTCTAATGTTGCTGTAGCAAGGTCACAGGA 6252
    |||
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
    |||
6251 CGATATACGACAAATGATGTTTATGTTTATCCAGCTGATATTATTCGAAGATACGGGT 6192
    |||
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
    |||
6191 AATGCTTATATGTTCTCCTCATGGAGGTCATCATCACTACATTCCTCCAAAAGCGATTTATCT 6132
    |||
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg 218
    |||
6131 GCTAGTGAATTAAGCAGCAGCTAAAGCACAATCTGGCTGGAAAAAATATGCAACCGAGTCAG 6072
    |||
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
    |||
6071 TTAAGCTATTCTTCAACAGCTAGTACAAATAACACGCAA-----TCTGTAGCAAAA 6021
    |||
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
    |||
6020 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAAATCTCCAGAGTCTTTTG 5970
    |||
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
    |||
    |||
```


535	AlaTyrValThrProHisMetThrHisSerHisTyrPleLysLysAspSerLeuSerGlu	554	QY
3	GCCTATGTAACTCCACATATGACCCATAGCCATCGGATTAAAAAAGATAGTTGTCTGAA	62	Db
555	AlaGluArgAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThr	574	QY
63	GCTGAGAGCGGCACC - CAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCTCTCGACA	121	Db
575	AspHisGlnAspSerGlyAsnThrGluAlaLysGlyValaGluAlaIleTyrAsnArgVal	594	QY
122	GACCATCAGGATTCAGGAAATACTAGGACCAAGGACGAGACTATCTACAACCCCGTG	181	Db
595	LysAlaLalysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu	614	QY
182	AAAGCAGCTAAGAAGGTGCACCTTGATCGTATGCTTACAATCTTCAATATCTGTAGAA	241	Db
615	ValLysAsnGlySerLeuIleLeProHisTyrAspHisTyrHisAsnIleLysPheGlu	634	QY
242	GTCAAAAACGGTAGTTTAAATCACTACCTCATTTATGACCATTACCATTAACATCAAAATTGAG	301	Db
635	TripPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAla	654	QY
302	TGGTTTCAGCAAGGCCCTTTATGAGGCACCTTAAGGGGTATACTCTTGAGATCTTTGGCG	361	Db
655	ThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGly	674	QY
362	ACTGTCAAGTACTATGTTCGAACATCCAAACGAAGTCCGATTCAGATTAATGGTTTGGT	421	Db
675	AsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGlu	694	QY
422	AACGCTACGCACCATGTTCAAAGAAA CAAAAATGGTCAAGCTGATACCAATCAACCGAA	481	Db
695	LysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArgGluGlu	714	QY
482	AAACCAAGCAGGAGAAACCTCAGACAGAAATAA CCTGAGGAAGAAACCCCTCGAGAAGAG	541	Db
715	LysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSerPro	734	QY
542	AAACCGCAAGCGAGAACACAGAGTCTCCAAACCACAGAGAAC CAGAGAATCACCA	601	Db
735	GluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGluAla	754	QY
602	GAGGAATCAGAAGAACCTCAGGTCGAGACTGAAAGGTTGAGAAAAA CTGAGAGAGCCT	661	Db
755	GluAspLeuLeuGlyLysIleGlnAsp	763	QY
662	GAGATTACTTCGAAAAATATCCAGGAT	688	Db

RESULT 13
 JS-09-468-656A-7
 Sequence 7, Application US/09468656A
 Patent No. 6582706
 GENERAL INFORMATION:
 APPLICANT: Johnson, Leslie S.
 APPLICANT: Adamou, John E.
 TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
 TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural
 TITLE OF INVENTION: Motifs
 FILE REFERENCE: 469201-444
 CURRENT APPLICATION NUMBER: US/09/468,656A
 CURRENT FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: 60/113,048
 PRIOR FILING DATE: 1998-12-21
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 7
 LENGTH: 1455
 TYPE: DNA
 ORGANISM: Streptococcus pneumoniae
 JS-09-468-656A-7

Alignment Scores:	8.48e-102	Length:	1455
pred. No.:			

Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best local Similarity: 45.52% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 4 Gaps: 14

US-09-765-272A-66 (1-763) x US-09-468-656A-7 (1-1455)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 61 TGTGCTTATGCACCTAACCAACGATCGTTCG---CAGGAAATTAAGACAAATATCTGTGTC 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGTGGATGCGAGCCAGTCAAGTCAGAAAAGTGAAAACCTTGACACCAGACGAGTT 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAAGGAATTCAGGCTGAGCAAAATGTGAATCAAAATTACAGATCAGGCTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTAACGTACACGGTGACCACTATCAATTACTATTAATGGAAGATTCCTTATGATGCCCT 297
QY 81 IleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 TTTAGTCAAGAACTCTTGATGAAGGATCCAAACTATCAACTTAAGACGCTGATTTGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 358 AATGAAGTCAAGGGTGGTTATATCATCAAGTCGATGGAAATATATATGCTTACCTGAAA 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGCAGCTCATGCTGATATATGTTCCAACTAAAGATGAATCAATCGTCAAAAAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly 160
Db 478 CATGTCAAAGAT--AATGAGAAGGTTAACTCTAATGTGTTGTAGCAAGGCTCAGGGA 534
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACGCAAAATGATGGTTATGTCTTTAATCCAGCTGATATATATCAAGATACGGGT 594
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTCTCTCATGGAGGTCACATCATCATCATTCCTCCAAAGACGATTA 654
QY 201 AlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg 218
Db 655 GCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAG 714
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGCTATTCTTCAACAGCTAGTGACAAATAACACGCAA-----TCTGTAGCAAAA 765
QY 239 AsnLeuThrValThrProThrTyrHisGluAsnGlnGlyGluAsnIleSerSerLeuLeu 258
Db 766 GGATCAACTAGCAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG 816
QY 259 ArgGluLeuTyrAlaLysProLysSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 817 AAGGAACTCTATGATTCACCTAGCGCCCAACGTTACAGTGAATCAGATGGCTGCTGCTTT 876
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCCCTGCTAAGATTAATCAGTCGTACACCAATGGAGTTGGATTCGCATCGCGACCAT 936
QY 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle 318
Db 937 TACCATTATTCCTTACAGCAGCTTTCGCTCTAGAGAAAAGATTGCCAAGATGGTG 996
QY 319 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338

```
Db 997 CCTATCAGTGAAGTGGTCTTACAGCTTTCTACAAATGCAGAACTTAAT----- 1044
Qy 339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAGTAGTGTCTAGTCTAGGCGAGTCTTCA 1074
Qy 359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCTCTCTTTTACGACAGTAGAGG---CTCTCTTCAGCATCTGATGGTTAT 1131
Qy 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 ATTTTAAAT-----CCAAAAGATATCGTTCAGAGAAACG 1164
Qy 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysSerHisLysLeuGlyAla 418
Db 1165 GCTACA-----GCTTATATTAGACATGCTGATCATTTCCATTAC----- 1206
Qy 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAATCAAAT----- 1221
Qy 439 AlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAAATTTGGCAACCGACTCTTCCAAACAATAGCTAGCAACACCTTCTCCATCTCT 1278
Qy 459 AspAsnLeuLeuGluArgLeuLysAspVal**SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
Qy 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CAATC-----AATCCAGGA 1293
Qy 499 IleThrTyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
Qy 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GATGGATACGGATTGTATGCTTAATGCTTATTCGCTGAAGATGAATCAGGTTTGTCTATG 1377
Qy 539 ProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGlu 556
Db 1378 ACTCAGGAGACCAATCATATTTCTTCAGAGAGGACTTGACAGAGACGA 1431

RESULT 14
US-08-961-083-181
; Sequence 181, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 181:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-181

Alignment Scores:
Pred. No.: 2,26e-99 Length: 1342
Score: 1141.50 Matches: 250
Percent Similarity: 59.41% Conservative: 75
Best Local Similarity: 45.70% Mismatches: 119
Query Match: 28.40% Indels: 103
DB: 3 Gaps: 14

US-09-765-272A-66 (1-763) x US-08-961-083-181 (1-1342)
Qy 5 LeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSerTyrIleAsp 24
Db 2 CTAACCCAGCATCGTTCG---CAGGAAATATAGGACAATATCGTGTCTCTTATGTGGAT 58
Qy 25 GlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSerLysArgGlu 44
Db 59 GGCAGCCATCGATGTCAGAAAGTGAAGAACTTCACACCAGCAGGTTAGCCAGAAAGAA 118
Qy 45 GlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyrValThrSerHis 64
Db 119 CGAATTCAGGCTGAGCAAAATTGTAATCAAAATTACAGATCAGGCTATGTAAAGTCACAC 178
Qy 65 GlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIleIleSerGluGlu 84
Db 179 GGTGACCATCTATCTACTATATATATGGAAGTTCTTATGATGCCCTCTTTAGTGAAGAA 238
Qy 85 LeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsnGluIleLys 104
Db 239 CTCCTGATGAGGATCCAAACTATCACTTAAGAGCTGATATTGTCAATGAGTCAAG 298
Qy 105 GlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaHis 124
Db 299 GGTGGTTATATCATCAAGGTCGATGCAAAATATATATGCTCTACCTGAAAGATGCAGCTCAT 358
Qy 125 AlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGluArgSerHisAsn 144
Db 359 GCTGATAATGTTGCAACTAAAGATGAATCAATCGTCAAAACAAGAACATGTCAAAGAT 418
Qy 145 HisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGlyArgTyrThr 164
Db 419 ---AATGAGAGGTTAACTTAATGTTGCTGTAGCAAGGTTCTCAGGACCATATACACA 475
Qy 165 AspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIle 184
Db 476 AATGATGGTTATGCTCTTTTAAATCCAGCTGATATATTCGAGATACGGGTATGCTTATATC 535
Qy 185 ValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeu 204
Db 536 GTTCTCTCATGGAGGTCACTATCACTATCCAAAGGCAATTTATCTGCTAGTGAATTA 595
Qy 205 AlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArgProSerSer 222
Db 596 GCAGCAGCTAAAGCACATCTGCTGGAAAAAATATGCAACCGAGTCAGTTAAGCTATTCT 655
Qy 223 SerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrVal 242
Db 656 TCAACAGCTAGTGACAAATAACACGAA-----TCTGTAGCAAAAGGATCACTAGC 706
```



```
Db 839 TTGGCGACTGTCAAGTACTATGTCGAACATCCAAACCAAGCTCG-CATTGAGATAATGGT 781
Qy 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
Db 780 TTGTGTAACGCTAGCGACCATGTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAA 721
Qy 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArg 712
Db 720 ACGGAAAAACCAACAGAGGAAACCTCGACAGAAAAACCTGAGGAGAAACCCCTCGA 661
Qy 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu----- 729
Db 660 GAAGAGAAACCGCAAAAGCGAGAAACAGAGTCTCCAAAACCAACAGAGGAAACCAAGAA 601
Qy 730 -----ProGluLysProGluGluSerGluGluProGlnValGluThrGluLysVal 747
Db 600 GAATCAACAGAGAAATCACAGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAGGTTA 541
Qy 748 GluGluLysLeuArgGluAlaGluAspLeuLysGlyLysIleGlnAsp 763
Db 540 ---AGAAACTGAGAGGCTGAGATTACTTGGAAAAATCCAGAT 496

RESULT 17
US-08-526-840B-34/c
; Sequence 34, Application US/08526840B
; Patent No. 6001564
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
; NUMBER OF SEQUENCES: 177
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,840B
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/304,732
; FILING DATE: 12-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 841 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM SOURCE: Streptococcus pneumoniae
; ORGANISM: Streptococcus pneumoniae
US-08-526-840B-34
```

```
Alignment Scores:
Pred. No.: 1.36e-39 Length: 841
Score: 512.00 Matches: 104
Percent Similarity: 93.10% Conservative: 4
Best Local Similarity: 89.66% Mismatches: 2
Query Match: 12.74% Indels: 7
DB: 3 Gaps: 2

US-09-765-272A-66 (1-763) x US-08-526-840B-34 (1-841)
Qy 653 LeuAlaThrValLysThrValGluHisProAsnGluArgProHisSerAspAsnGly 672
Db 839 TTGGCGACTGTCAAGTACTATGTCGAACATCCAAACCAAGCTCG-CATTGAGATAATGGT 781
Qy 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
Db 780 TTGTGTAACGCTAGCGACCATGTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAA 721
Qy 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArg 712
Db 720 ACGGAAAAACCAACAGAGGAAACCTCGACAGAAAAACCTGAGGAGAAACCCCTCGA 661
Qy 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu----- 729
Db 660 GAAGAGAAACCGCAAAAGCGAGAAACAGAGTCTCCAAAACCAACAGAGGAAACCAAGAA 601
Qy 730 -----ProGluLysProGluGluSerGluGluProGlnValGluThrGluLysVal 747
Db 600 GAATCAACAGAGAAATCACAGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAGGTTA 541
Qy 748 GluGluLysLeuArgGluAlaGluAspLeuLysGlyLysIleGlnAsp 763
Db 540 ---AGAAACTGAGAGGCTGAGATTACTTGGAAAAATCCAGAT 496

RESULT 18
5231168-1
; Patent No. 5231168
; APPLICANT: DIEGIEL, MORTEN, BORRE, MARTIN; JEPSEN, SOREN;
; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 1
; LENGTH: 3095
5231168-1

Alignment Scores:
Pred. No.: 1.05e-06 Length: 3095
Score: 177.00 Matches: 136
Percent Similarity: 36.19% Conservative: 109
Best Local Similarity: 20.09% Mismatches: 234
Query Match: 4.40% Indels: 198
DB: 6 Gaps: 34

US-09-765-272A-66 (1-763) x 5231168-1 (1-3095)
Qy 177 GluAspThrGlyAspAlaThrIleValProHisGly-----AspHisTyr 191
Db 103 GAAGAATCAGGTGAAAGTAAATAGTTGATAATGATGAAGTGGTGGTTTGAAGAGAGCTCAT 162
Qy 192 HisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrp 211
Db 163 CATGAAAAATTTTCATCTCGAAGTAAAGTAACTCTGAATTAATAATCAAAATCAATTTGTGAA 222
Qy 212 AsnGlyLysGlnGlySerArgProSerSer-----SerSerSerTyrAsn 226
Db 223 TCTGACAAAAGTGTAACTGAACTCTGAAATGAAAGAGTGTGTATCTGAAGAAAGCAAC 282
Qy 227 AlaAsnProAla-----GlnProArgLeuSerGluAsnHisAsnLeuThrValThr 243
Db 283 CTTGAAACCGCTGAAAAATGAAGAAAGTAGTAGAAGAGCTCATCAGGAGAAAGTAATGTA 342
```



```
Db 1603 ATGATGAAATTTTG-----ATAAT-CATCTCGAATAATATTTTAAAGAAAGATATTA 1655
QY 498 nileThrTyThrAspAsp---GluileGlnValAlaLysLeuAlaGlyLysTyThrTh 517
Db 1656 TTATGATGAAATGATGATGAGTAAAGATTAAATAATAGGT-----GTCACATT 1709
QY 517 rGluAspGlyTyIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyVa 537
Db 1710 AAAAAA-----TTTGAACA----- 1725
QY 537 lThrProHisMetThrHisSerHisThrIleLysLysAspSerLeuSerGluAlaGluAr 557
Db 1726 -----CTTAAATGGAATCTTATGTAACCATTA 1757
QY 557 gAlaAlaAlaGlnAlaTyAlaLysGluLysGlyLeuThrProSerThrAspHisGl 577
Db 1758 ATTGATT----- 1764
QY 577 nAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyAsnArgValLysAlaAl 597
Db 1765 -CATTTAGGAATAAAGATAAAACACATTGGAAGCTATAAACACGATATTCAAATTAT 1823
QY 597 aLysLysValProLeuAspArgMetProTyAsn-----LeuGlnTyThrValGluVa 615
Db 1824 TAAACRA---CAATTACAAGCTATT---TATAATGAACTTATGAATTATACA----- 1869
QY 615 LysAsnGlySerLeuLeuIleProHisTyAspHisTyAsnIleLysPheGluTr 635
Db 1870 ---AATGGAAC-----AAAATATTCAA---CAAAT 1895
QY 635 pPheAspGluGlyLeuTyGluAlaProLysGlyTyThrLeuGluAspLeuLeuAlaTh 655
Db 1896 ATTTCAACAAATATTCTAGAA-----AATGATGTTCTT----- 1929
QY 655 rValLysTyThrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAs 675
Db 1930 -----AA 1931
QY 675 nAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLy 695
Db 1932 TCAAGAACCGAGGAGAAATGGAATAACAAGTTGAGCAATCACCAGCAATAGAGAAGC 1991
QY 695 sProSerGluLysProGlnThrGluLysProGluGluThrProArgGlu----- 713
Db 1992 TGAAGTCGATGCCCTCGCACCAAAAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2051
QY 714 ---GluLysProGlnSerGluLysProGlnSerProLysProThrGluGluProGluGl 732
Db 2052 AAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2111
QY 732 uSerProGluLysSerGluGluProGlnValGluThrGluLysValGluGluLysLeuAr 752
Db 2112 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2171
QY 752 gGluAlaGluAspLeuLeu 758
Db 2172 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2190
```

RESULT 20

```
US-09-216-393B-331/c
; Sequence 331, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 331
```

```
; LENGTH: 2392
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
US-09-216-393B-331
Alignment Scores:
Pred. No.: 9,77e-06 Length: 2392
Score: 165.00 Matches: 133
Percent Similarity: 34.07% Conservative: 84
Best Local Similarity: 20.88% Mismatches: 254
Query Match: 4.11% Indels: 166
DB: 4 Gaps: 29
US-09-765-272A-66 (1-763) x US-09-216-393B-331 (1-2392)
QY 205 AlaAlaAlaGluAlaTyThrAsnGlyLysGlnGlySerArgProSerSerSerSer 224
Db 2336 TCAGCAGCGCTACTGTGTGGACCTCCAG-----ACAAGTTGAGTCATCACCTCGC 2283
QY 225 TyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrPro 244
Db 2282 AAGTCGGCTCCCTGCCCGCCCGACTGTGCCCCCGCGAGGGTGGTGAAGAGACG 2223
QY 245 ThrTyHisGlnAsnGlnGly-----GluAsnIleSerSerLeuLeuArgGluLeuTy 262
Db 2222 GCAACGCATGTGGACCGTGGTCCGTGGAGCCGTGTCCCGG---CGAAGGAAACAACAT 2166
QY 263 AlalysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282
Db 2165 GAGCACCCCGCGAGAGAGACACCGAGAG----- 2136
QY 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyHisPheIle 302
Db 2135 -----CGAGCGGTGCGTCAACCAAGTCGAAC-----CATGGAGTG 2100
QY 303 ProTyGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIlePro----- 319
Db 2099 CCC-----AAGTCAGG-CTGGGCAAGCATCGGCAACTATGTCGCTTCCGACT 2050
QY 320 -----LeuArgTyArgSerAsnHisTrpValProAspSerArgProGluGlnProSer 337
Db 2049 GTGGAATGCTCTCGCGCTTCGCCATGG-----CCTTCC 2014
QY 338 ProGlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAla 357
Db 2013 CCGAGAACTGCCAGAGCTCGTGGCTTCGCG-----CCTGC-GATGCT 1970
QY 358 ProSerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly 377
Db 1969 CCGCGCTCCCGCAAGAGGACCGCTGC----- 1943
QY 378 TyrValPheGluGluAsnGlyValSerArgTyThrIleProAlaLysAspLeuSerAlaGlu 397
Db 1942 CATCTTTCTCGGCTGGTCCCAAGTGCACACATACCCCGGCACTACTCTGTACGAGCAG 1883
QY 398 ThrAlaAla-----GlyIleAspSerLysLeuAlaLysGlnGluSerHisLys 415
Db 1882 AGCGCTCTCGCATGGCATGATTGACCGAGTCCCGCTTCTGCACTCCCGAGGAG 1823
QY 416 LeuGlyAlaLys---LysThrAspLeuProSer-----SerAspArgGluPheTy 431
Db 1822 GTCGGCTCGGAGCTTTCACCTGAGTCGCTTCGAATGCGGTTCCCTCGCGAGTTCGGC 1763
QY 432 AsnLysAlaTyAspLeuLeuAlaArgIleHisGlnAspLeuAspAsnLysGly--- 450
Db 1762 GAGTGTGTGAACGGCTTCAGGAGAGAGAGTACTCGGACTGCCCGCCGATCATAGAAGATC 1703
QY 451 ArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***Ser 470
Db 1702 CGTCAGTGTCTGACGAATCTCTGCTCTGCTTCCGCGAG----- 1664
QY 471 AspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIleArgHisProGlu 490
```


512 ACSCGCTCTCGCATGGCATGATTGACCGAGTCCGCTTCTGCATCTCCGACGAGAG 571
416 LeuGlyAlaLys---LysThrAspLeuProSer-----SerAspArgGluPheTyr 431
572 GTCGGCTCGGACGCTTCCATCGAGCTTCCGAATGCGGTTCTCCCTCGCGGAGTTCGGC 631
432 AsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAsnLysGly--- 450
632 GAGTGTGTGAACCGCTTCAGGAGAGAGCTACTCGGACTGCCCGCATCAAGAAGATC 691
451 ArgGlnValAspPheGluAlaLeuAsnLeuLeuGluArgLeuLysAspVal***Ser 470
692 CGTCAGTGTCTGACGAATCTGCTCTGCTTCGCTTCGCGCGAG----- 730
471 AspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIleArgHisProGlu 490
731 -----TGTCACCTCGCGGGAACCCAG 754
491 -----ArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAsp 503
755 CAAGGCTCGGTATCCGCAAGAGAGCTGCATCGCAACAGCTGCTGCGCTGTGTGCGAG 814
504 AspGluIle-----GlnValAlaLysLeuAlaGlyLys---TyrThr 516
815 GCGGAGGTCTGCGCGGATGTCACCCAGAGATTGAGGAGGAAGAGCAACATTTCCCC 874
517 ThrGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAsp----- 534
875 CCGTAGAGAGCGAGGCTTGCCTCCATATGAAGAGGGTCTGCTGAGGTGAGCTTGT 934
534 ----- 534
935 CTTCCCGAGGAGGAGATCCTGAAGAGAAACATGTTCTGAGGAGGAAATCCTCGAAGGA 994
535 AlaTyrValThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGlu 554
995 GAACATATTCCTGAAGAGCTCCCAAGAGGCGAGCATGTTCTGAGGAGGAAATCCTGAA 1054
555 AlaGluArgAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThr 574
1055 GGAGACATGTTCTGAGAGGAATC-----CCTGAAGGA 1090
575 AspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVal 594
1091 GAGCATGTTCTGAAGAGATCCCGAGAA-----GGCGAACATGTTCTGAGGAGGAAATC 1144
595 LysAlaAlaLysLysValProLysAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614
1145 CCGTAAGAGGAACATGTTCTGAGGAGGAATCCTGAAGGAGAACATGTTCTCTGAAGAG 1204
615 ValLysAsnGlySerLeuIle-----IleProHisTyrAspHisTyrHisAsnIle 631
1205 TTCCCTGAAGGAGAACATATCTCTGAGGAGCTCCTGAAGGAGGACATATTCAT----- 1258
632 LysPheGluTrpPheAspGluGlyLeuTyr-----GluAlaProLysGly---Tyr 647
1259 -----GAGAGTTCCTGAAGGAGAGCATATTCATGAGGAGCTCCTGAAGGCGAGCAT 1312
648 ThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgPro 667
1313 GTTCCTGAAGAGGAGATCCTCGAAGGAGAACATATTCCTGAGGAGTTCCTCGAGGCGAG 1372
668 HisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGln 687
1373 CATGTTCTGAGGAGGAATCCTCGAAGGAGAACATATTCCTGAGGAGGAGTTCCTCGAA 1432
688 AlaAspThrAsnGlnThrGluLys-----ProSerGluLys 700
1433 GGAGAGCATGTTCTGAGGAGGAGATCCTCGAAGGCGAGCATGTTCTGAGGAGGAGCTC 1492
701 ProGlnThrGluLys---ProGluGluGluThrProArg----- 712
1493 CCGTAGGAGGAACCTTATTCCTGAGGAGGAGATCCTCGAAGGAGGAGCATGTTCTCGTAGAGAG 1552

QY 713 -----GluGluLysProGlnSerGluLysProGluSerProLysProThrGlu 728
Db 1553 CTCCTCGAAGCGAGCATGTTCTGAGGAGGAGATCCTCGAAGGAGCATGTTCTCTGAA 1612
QY 729 Glu-----ProGluGluSerProGluGluSerProGluGlu 739
Db 1613 GAGGAAATCCTCGAAGCGGAGCATGTTCTGAGGAGGAGACCCCTGAAGGAGAACATGCT 1672
QY 740 ProGlnValGluThrGluLysValGluLysLeuArgGluAlaGluAsp 756
Db 1673 CCAGAGGAAGAGATCTCTGCACCTGAGGAGACCGAAAGAGGAGGAAGAA 1723

RESULT 22

US-08-242-932-1
; Sequence 1, Application US/08242932
; Patent No. 5595740
; GENERAL INFORMATION:
; APPLICANT: Brady, L. Jeanmine
; TITLE OF INVENTION: Cloning of No. 5595740-Iga Fc Binding Forms of
; TITLE OF INVENTION: the Group B Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,932
; FILING DATE: 16-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UFI42
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-242-932-1

Alignment Scores:

Pred. No.:	3,73e-05	Length:	4200
Scores:	163.00	Matches:	166
Percent Similarity:	35.26%	Conservative:	127
Best Local Similarity:	19.98%	Mismatches:	316
Query Match:	4.06%	Indels:	222
DB:	1	Gaps:	44

US-09-765-272A-66 (1-763) x US-08-242-932-1 (1-4200)

QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25
Db 917 CATCAAAAAGTCAGTTGAAATAATGCGAGAGCAAAAGGGAATCACAATCGAAGATAAA 976
QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35
Db 977 GATTCTATGCTGAAAAAATCGAAGATATTCGTAACAAGCTCAACAGCAGATAAAAA 1036
QY 36 ThrProAspGluValSerLysArgGlu----- 44

QY 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721
 Db 2858 GCTCCAGACACACCGCATGTTCCGGAATCACCAAGGCCCCAGAACACCGCGTGTCCG 2917
 QY 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739
 Db 2918 GAATCAACCAAGACTCCAGAGACGCCATGTTCCGGAATCACCAAGGCCCCAGAACCA 2977
 QY 740 ProGlnVal---GluThrGluLysValGluGlu 749
 Db 2978 CCGGTGTTCCGGAATCACCAAGACTCCAGAA 3010

RESULT 23

US-08-714-481-1
 ; Sequence 1, Application US/08714481
 ; Patent No. 5766606
 ; GENERAL INFORMATION:
 ; APPLICANT: Brady, L. Jeannine
 ; TITLE OF INVENTION: Cloning of No. 5766606-IGA Fc Binding Forms of
 ; TITLE OF INVENTION: the Group B Streptococcal Beta Antigens
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Saliwanchik & Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: FL
 ; COUNTRY: USA
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/714,481
 ; FILING DATE: 16-SEP-1996
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/242,932
 ; FILING DATE: 16-MAY-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UF142
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; TELEFAX: 904-372-5800
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4200 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-714-481-1

Alignment Scores:
 Pred. No.: 3,73e-05 Length: 4200
 Score: 163.00 Matches: 166
 Percent Similarity: 35.26% Conservative: 127
 Best Local Similarity: 19.98% Mismatches: 316
 Query Match: 4.06% Indels: 222
 DB: 1 Gaps: 44

US-09-765-272A-66 (1-763) x US-08-714-481-1 (1-4200)

QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25
 Db 917 CATCAAAAAAGTCAAGTTGAAAAAATGCGACAGCAAAAGGAATTCACAAATGAAGATAAA 976
 QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35

Db 977 GATTCCTAGCTGAAAAAATCGAAGATATTCGTAAACAAAGCTCAACAAAGCAGATAAAAA 1036
 QY 36 ThrProAspGluValSerLysArgGlu----- 44
 Db 1037 GAAGATGCCGAGTAAGGTTCTGTGAAGAACTAGGTAAACTCTTTAGTCAACTAAAGCT 1096
 QY 45 GlyIleAsnAlaGlu-----Gln***ValIleLysIleThrAspGlnGlyTyrValThr 62
 Db 1097 GGTCTGGATCAAGAAATTCAGAGCATGTGAAGAAACGAGTGTAGTACGAGAAATACT 1156
 QY 63 SerHisGlyAspHisTyrHisTyrAsn-----GlyLys 74
 Db 1157 CAGAAAGTTGAGAA---CACTATGCTAATACGCTTCAGAACCTTCTCTCAAAAATCTCTT 1213
 QY 75 ValProTyrAspAlaIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeu 94
 Db 1214 GAAGAACTAGATAAGGCACTACCAATGAACAGCTACACAGTTAAATAATCAATCTTA 1273
 QY 95 LysAspSerAspIleValAsnGluLysGlyTyrValIleLysValAsnGlyLys 114
 Db 1274 GAAAAAGCTCAAAAGCTCAAGAAATACAA---CCTCTTATCAAGAAACGAATGTGAAA 1330
 QY 115 TyrTyrValTyrLeuLysAspAlaIleHisAlaAspAsnIleArgThrLysGluGluIle 134
 Db 1331 TTGTATAAGGCTATG-----AGTGAGAGCTTG 1357
 QY 135 LysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAla 154
 Db 1358 GAGCAGGTTGAGAGGAATTAACAATATTCGGAAGCTTAATTTAGAGATTTGGTTGCG 1417
 QY 155 AlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAsp 174
 Db 1418 AAATCTAAA-----GAA 1429
 QY 175 IleIleGluAspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIle 194
 Db 1430 ATCGTAAGAGAA-----TACGAAGGAAAA 1453
 QY 195 ProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLys 214
 Db 1454 CTTAATCAATCTAAAAATCTTCAGAAATTAAGCAACTAGAGAGAGAAAGCTCAATTCGAG 1513
 QY 215 -----GlnGlySerArgProSerSerSerSerSerSerSerSerSerSerSerSer 229
 Db 1514 TTGAACAAGTTGTGGAGGATTTTAAAAAATTTAAACGTCAGAGCAAGTGACACCA 1573
 QY 230 -----AlaGlnProArgLeuSerGluAsnHisAsn-----Leu 240
 Db 1574 AAAAAACGTGTCAACAGAGATTTAGCTGCTAATGAAAAATATCAACAAAAAGATTGAGTTA 1633
 QY 241 ThrValThrProThrTyrHisGlnAsn-----GlnGlyGluAsnIleSerSer 256
 Db 1634 ACAGTTTCACCA-----GAGAAATATCATGTATATGAAGGTGAAGCGGAAATTT 1684
 QY 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276
 Db 1685 ACAGTCACA-----GCTAAAAAGTGTTCGAGACCGCTGGACCTTCAGTGATCTT 1735
 QY 277 Ile-----PheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValPro 294
 Db 1736 TTAACAAAATATAATCCGTCT---GTATCAGATAGAAATTAGTACAAATATATAAGACTAAC 1792
 QY 295 HisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIle 314
 Db 1793 ACGGATAATCATAG-----ATTCCGAAAAATCACTATCAAGAAATTTGAAGCTAAATGAA 1846
 QY 315 AlaAlaIleIleProLeuArgTyrArgSerAsn-----His 326
 Db 1847 AGTCAACACAGTGACTCTAAAAGCTAAAGATGATTTCTGGCAATGTAGTTGAAAAAACATTTC 1906
 QY 327 TrpValProAspSerArgProGlnGlnProSerProGlnSerThrProGluProSerPro 346
 Db 1907 ACTATTACAGTGCAAAAGAGAGAGAGAAACAAGTTCTTAAACACACAGCAGAGAAAGAT 1966

347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLeu 366
1967 TCTAAACCGAAGAAAGGTTCTCTCAAGAACCAAAATCAAT-----GACAAATCAA 2020
367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
2021 TTACAAGAGTGTGATTAAATCAGCTCAACAGAACTCGAAAAGTTAGAAAAGCAATAAA 2080
387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
2081 GAATTAATCGGACCAACAGAGATTCATCCATCCAGAGTATGTTTCAAAAATCTATT 2140
407 -----AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuPro 424
2141 TGGGAGTCAAAAGAGCTCTCAGGAAGGCATACAAAGTTTAAAGAGATTATTGGT 2200
425 SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAlaArgIleHisGlnAsp 444
2201 GATTCATCTTCAAAATACACAGAGCACTAT----- 2233
445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuArg 464
2234 -----TTTAAACAATATAAATCT--GATTT-----ATGAATTATCACTTCATGCA 2278
465 LeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAla 484
2279 CAAATGGAGATCTGACTAGAAAAGTGGTTCAGTATATGAAC----- 2320
485 ProIleArgHisProGluArgLeuGlyLysProAsnAlaGluIleThrTyrThrAspAsp 504
2321 -----AAATATCCCTGAT-----AATGCAGAAATATAA----- 2347
505 GluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAsp 524
2348 -----AGATATTGAGTCAGATATGAGAGACGAAAGAGATTAATAC----- 2392
525 ProArgAspIleThrSerAspGluGlyAspAla----- 2392
2393 -----GGAAGTTTAGAAATGATGCTTTGAAAGGCTATTTTCAGAAATATTTC 2440
537 ValThrProHisMetThrHisSerHisThrLysLysAspSerLeuSerGluAlaGlu 556
2441 CTTACACCATTTAATAA-----ATTAGCAGATGTTAGATGATTGGAT 2485
557 ArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHis 576
2486 AAAAAAGTAGAACAA-----GATCAGCCAGCACCATTCCGAAAATTCAGAA 2533
577 GlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAla 596
2534 ATGGATCAGGCTAAGGAAAGGCTAAGATTGCTGTATCGAAGTAT----- 2578
597 AlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLys 616
2579 ATGAGTAAGGTT--TTAGATGGAGTTTCAATCAATCTCAG-----AAGAAA 2623
617 AsnGlySerLeuIleProHisTyrAspHisTyrHisAsnLysPheGluTyrPhe 636
2624 AATAACAGTAATTTGTT-----GATCTTTT 2650
637 AspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrVal 656
2651 AAGGAA-----CTTGAAGCGATTAAACAACTATTTTGTAT----- 2689
657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsn-----Gly 672
2690 -----ATTGACAATGCAAGACTGAAGTAGAGATTGATACTTAGTACAGATGCA 2740
673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
2741 TTCTCAAAATGAATGCTACTGTTCTGTAATTTTCAAAAGGT-----CTAGAGACAAATACG 2797

QY 693 ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708
Db 2798 CCAAGAACTCCAGATACACCGAAGATTCCAGAGCTACCTCAAGCCCGCATACACGCGAG 2857
QY 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721
Db 2858 GCTCCAGACACACCGCATGTTCCGGAATCACCAGAGCCCGCGGTGTTCCG 2917
QY 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739
Db 2918 GAATCACCAGAAAGACTCCAGAGACCGCATGTTCCGGAATCACCAGAGCCCGCATACGCA 2977
QY 740 ProGlnVal---GluThrGluLysValGluGlu 749
Db 2978 CCGCGTGTCCGGAATCACCAGAGACTCCAGAA 3010
RESULT 24
US-08-923-992A-1
; Sequence 1, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 320..3811
US-08-923-992A-1

Alignment Scores:
Pred. No.: 3,73e-05 Length: 4200
Score: 163.00 Matches: 166
Percent Similarity: 35.26% Conservative: 127
Best Local Similarity: 19.98% Mismatches: 316
Query Match: 4.06% Indels: 222
DB: 3 Gaps: 44

US-09-765-272A-66 (1-763) x US-08-923-992A-1 (1-4200)

2690 -----ATTGCAATGCAAGACTGAAGTAGAGATTGATACTTAGTACACGATGCA 2740
|||||
b PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
|||||
y TTTCTCAAAATGAATGCTACTGTTGCTAAATTTCAAAAGGT---CTAGAGACAATACG 2797
|||||
b ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708
|||||
y CCAGAACTCCAGATACACCGAAGATTCCAGAGCTACCTCAAGCCCCAGATACACCGCAG 2857
|||||
b GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721
|||||
y GCTCCAGACACACCGCAGTTTCCGGAATCATCCAAAGGCCCCAGAGCACCGCGTGTCCG 2917
|||||
b GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739
|||||
y GAATCACCAAGACTCCAGAGCACCGCATGTTCCGGAATCACCAAGGCCCCAGAGCA 2977
|||||
b ProGlnVal---GluThrGluLysValGluGlu 749
|||||
y CCGCGTGTTCGGAATCACCAAGACTCCAGAA 3010
|||||
b

ESUL 25

CT-US95-06111-1

Sequence 1, Application PC/TUS9506111

GENERAL INFORMATION:

APPLICANT: Street address: 186 Grinter Hall
APPLICANT: City: Gainesville
APPLICANT: State/Province: Florida
APPLICANT: Country: US
APPLICANT: Postal code/Zip: 32611
APPLICANT: Phone number: 904-392-8929
APPLICANT: Fax number: 904-392-6600
APPLICANT: Telex number:
TITLE OF INVENTION: Cloning of Non-IgA Fc Binding Forms of
NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06111
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/242,932
FILING DATE: 16-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFI42
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4200 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

CT-US95-06111-1

Alignment Scores: 3.73e-05 Length: 4200
Pred. No.: 163.00 Matches: 166
Score: 35.26% Conservative: 127
Percent Similarity: 19.98% Mismatches: 316
Best Local Similarity: 4.06% Indels: 222
Query Match: 5 Gaps: 44
DB: 44

US-09-765-272A-66 (1-763) x PCT-US95-06111-1 (1-4200)

QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25
|||||
Db 917 CATCAAAAAAGTCAAGTTCGAAAAATGCGCAGACGCAAAAGGGAATCACAAAATGAAGATAAA 976
|||||
QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35
|||||
Db 977 GATTCTATCTGCTGAAAAAATCGAAGATATTCGTAACCAAGCTCAACAGCAGATAAAAA 1036
|||||
QY 36 ThrProAspGluValSerLysArgGlu----- 44
|||||
Db 1037 GAAGATGCCGAAGTAAGGTTCGTGAAGAACTAGGTAAACTCTTTAGTTCACCTAAAGCT 1096
|||||
QY 45 GlyIleAsnAlaGlu-----Gln***ValIleLysIleThrAspGlnGlyTyrValThr 62
|||||
Db 1097 GGTCTGGATCAAGAAATTCGAAGCATGTGCAAGAAAGAAACGAGTAGTGAGGAAATACT 1156
|||||
QY 63 SerHisGlyAspHisTyrHisTyrAsn-----GlyLys 74
|||||
Db 1157 CAGAAAGTTGTATGAA---CACTATGCTAATAGCTTCAGAACCTTGCTCAAAAATCTCTT 1213
|||||
QY 75 ValProTyrAspAlaIleIleSerGluGluLeuMetLysAspProAsnTyrGlnLeu 94
|||||
Db 1214 GAAGAACTAGATAAGCAACTACCAATGAACAAGCTACACAGCTAATAAAATCAATCTTTA 1273
|||||
QY 95 LysAspSerAspIleValAsnGluIleLysGlyTyrValIleLysValAsnGlyLys 114
|||||
Db 1274 GAAACGCTCAAAAGCTCAAGAAATACAA---CCTCTTATCAAGAAACCAAGTGTGAA 1330
|||||
QY 115 TyrTyrValTyrLeuLysAspAlaIleHisAlaAspAsnIleArgThrLysGluGluIle 134
|||||
Db 1331 TTGCTATAGGCTATG-----AGTGAGAGCTTG 1357
|||||
QY 135 LysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAla 154
|||||
Db 1358 GAGCAGTTGAGAGGAATTTAAACATAATTCGGAAGCTAATTTAGAGATTTGCTGCG 1417
|||||
QY 155 AlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyrIlePheAsnAlaSerAsp 174
|||||
Db 1418 AAATCTAAA-----GAA 1429
|||||
QY 175 IleIleGluAspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIle 194
|||||
Db 1430 ATCGTAAGAGAA-----TACGAAGAGAAA 1453
|||||
QY 195 ProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLys 214
|||||
Db 1454 CTTAATCAATCTAAAAATCTTCCAGAAATTAAGCAACTAGAGAGAGGAGCTCAATTCGAAG 1513
|||||
QY 215 -----GlnGlySerArgProSerSerSerSerSerSerSerSerSerSerSer 229
|||||
Db 1514 TTGAACAAGTTGTGGAGGATTTTAGAAAAAATTTAAACAGTCAAGCAGCAGTCAACCA 1573
|||||
QY 230 -----AlaGlnProArgLeuSerGluAsnHisAsn-----Leu 240
|||||
Db 1574 AAAAAACGTTCAACACGAGATTTAGCTGCTAATGAAATAATCAACAAAAAGATTGAGTTA 1633
|||||
QY 241 ThrValThrProThrTyrHisGlnAsn-----GlnGlyGluAsnIleSerSer 256
|||||
Db 1634 ACAGTTTCACCA-----GAGAATATCACTGTATATGAAGTGAAGCAGTGAATTT 1684
|||||
QY 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276
|||||

1685 ACAGTCACAC-----GCTAAAGATGATTGCAAGACGAGCTTGGACTTCAGTGATCTT 1735
277 Ile-----PheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValPro 294
1736 TTAACAAATATATCGTCT-----GATCAGATGAAATAGTACAAATTTATAAGACTAAC 1792
295 HisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIle 314
1793 ACGGATAATCATAG-----ATTGCCGAATCATCTATCAAGAATTTGAAGCTAAATGAA 1846
315 AlaArgIleIleProLeuArgTyrArgSerAsn-----His 326
1847 AGTCAAAACAGTACTCTAAAGCTTAAAGATGATCTCGCAATGTAGTTGAAAAACATTC 1906
327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346
1907 ACTATTACAGTGCAGAAAGAGGAGGAAACAAGTCTCTAAACACACAGACGAGAAAGAT 1966
347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
1967 TCTAAACCGAAGAAAGGTTCTCTCAAGACCAAAATCAAT-----GACAAAGTCAA 2020
367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
2021 TTACAAGAGTTGATTAAATCAGCTCAACAAGAACTGGAAAAGTTAGAAAAGCAATAAAA 2080
387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
2081 GAATTAATGGAGCAACGACAGATTCCTCAATCCATCCAGAGTATGTATTCAAAAATCTATT 2140
407 -----AlaLysGlnGlnSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuPro 424
2141 TGGGAGTCAAAAAGAGCGCTATCCAGGAAGCCATAACAGTTTAAAGAGATTATTTGGT 2200
425 SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAsp 444
2201 GATTCATCTTCAAAATACTACACAGACACTAT----- 2233
445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArg 464
2234 -----TTTAACAAATATAAATCT-----GATTTT-----ATGAATTATCACTTCATGCA 2278
465 LeuLysAspVal**SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAla 484
2279 CAATCGAGATGCTGACTAGAAAGTGGTTCAGTATATGAC----- 2320
485 ProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAsp 504
2321 -----AAATATCTGTAT-----AATGCAGAAATTAAT----- 2347
505 GluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAsp 524
2348 -----AAGATATTTGAGTCAGATATGAAGAGACGAAAGATAATATAC----- 2392
525 ProArgAspIleThrSerAspGluGlyAspAla-----Tyr 536
2393 -----GGAAGTTTAGAAATATGCTTTGAAAGCTATTTTGAGAAATATTTTC 2440
537 ValThrProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGlu 556
2441 CTTACACCAATTAATAA-----ATTAAGCAGATGTAGATGATTTGGAT 2485
557 ArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHis 576
2486 AAAAAAGTAGAACAA-----GATCAGCAGCAGCACCAATTCGGAAAAATTCAGAA 2533
577 GluAspSerGlyAsnThrGluAlaGlyAlaGluAlaIleTyrAsnArgValLysAla 596
2534 ATGCAGCTAGGCTAAGGAAAGCGCTAAGATGCTGTATCGAAGTAT----- 2578
597 AlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLys 616
2579 ATGAGTAAAGTT-----TTAGATGGAGTTCAATCAACATCTGCAG-----AAGAA 2623

QY 617 AsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPhe 636
Db 2624 AATAACAGTAAATGTT-----GATCTTTT 2650
QY 637 AspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrVal 656
Db 2651 AAGGAA-----CTTGAAGCGATTAAACAACAACAACTAATTTTTCAT----- 2689
QY 657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsn-----Gly 672
Db 2690 -----ATTGACAATGCAAAAGACTGAAGTAGAGATTGATAACTTAGTACAGATGCA 2740
QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
Db 2741 TTTCTAAAAAATGAATGCTACTGTCTCTAAATTTCAAAGGT-----CTAGACACAATACG 2797
QY 693 ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708
Db 2798 CCAGAAACTCCAGATACACCGAAGATTCCAGAGCTACCTCAAGCCCCCAGATACACCGCAG 2857
QY 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721
Db 2858 GCTCCAGACACACCGCATGTTCCTCGAATCACCAGAGGCCCCAGAACGCCGCGTGTTCG 2917
QY 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739
Db 2918 CAATCACCAGAGACTCCAGAGACCGCATGTTCCTCGAATCACCAGAGGCCCCAGAACGCC 2977
QY 740 ProGlnVal---GluThrGluLysValGluGlu 749
Db 2978 CCGCGTGTTCGGAATCACCAGAAAGCTCCAGAA 3010

RESULT 26

US-08-956-171E-39
Sequence 39, Application US/08956171E
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 4522 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 39:

JS-08-956-171E-39

Alignment Scores:

Pred. No.: 7,27e-05 Length: 4522

Score: 160.50 Matches: 175

Percent Similarity: 32.82% Conservative: 122

Mismatch: 19.34% Mismatches: 330

Query Match: 3.99% Indels: 279

Gaps: 40

JS-09-765-272A-66 (1-763) x US-08-956-171E-39 (1-4522)

```

2y 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21
bb 1220 AGCTATGAATGGACAGCATTTTGTGAATA-ATTAAGAGAGTACTGAAATATGAGC 1278
2y 22 TyrIleAsp-----GlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAsp 38
bb 1279 TGGTTGATAAATATTTCGCGAAGAT-----AATGATTCAAATGATGAC 1323
2y 39 GluValSerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGln 58
bb 1324 TTGATTCATAGAAGAAAAAGACGTCAGAATACACAAAATATAGATACGATCATGAC 1383
2y 59 GlyTyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAsp 78
bb 1384 TCATTACTGCCTCAAAATATGATATTATAGTCGCGAGGGGAAATTCGCTTTTCT 1443
2y 79 AlaIleIleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAsp 98
bb 1444 ATGACGTAGCTTATGAA-----AATCAAAATGTTGACAACTCT--- 1482
2y 99 IleValAsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyr 118
bb 1482 ----- 1482
2y 119 LeuLysAspAlaAlaHisAlaAspAsnIleArg---ThrLysGluGluIleLysArgGln 137
bb 1483 -----GCAGATACTATTTCAGATGAAAAAGAACAAATACCATCGAGAC 1524
2y 138 LysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArg 157
bb 1525 TATCGCAACAAAGCCAGATTCTCGTTCACAAAACGACAT-----CGC 1569
2y 158 AlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177
bb 1570 CGTAGAAGAAATCAAACTCAAGAA-----CAAAATATTATAGTGA 1611
2y 178 AspThrGlyAspAlaTyrIle-----ValProHisGlyAspHisTyrHisTyr 193
bb 1612 CAACGTGGGAATCTTAAAAATATACACGAAAGTATATAAATATCAATTCACATTC 1671
2y 194 -----IleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyr 210
bb 1672 CATCAATTAACCCAGGTACATATGTTTTCGCAATTT----- 1707
2y 211 TrpAsnGly-----LysGlnGlySerArgProSerSerSerSerTyrAsnAlaAsn 228
bb 1708 ---AATGGTATTGAGAGAAAGCCACACGCAACAAACACATATATGTTCTTAATAT 1764
2y 229 ProAlaGlnProArgLeuSerGluAsnHisAsnLeu---ThrValThrProThrTyrHis 247
bb 1765 -----ACAAATCATCGTGTAAAGATTCAACTCCAGATTATCAC 1803
2y 248 GlnAsnGln-----GlyGluAsnIleSerSerSerLeuLeuArgGluLeuTyrAlaLys 264

```

```

Db 1804 AAGAAAAGTTTCAAGACTTCAGAGGTACCGTCAGCTATTTTGGCACAATGAACCTAAA 1863
Qy 265 ProLeu-----SerGluArgHisValGluSerAsp 274
Db 1864 AAGTTAGAAAATGGTGTATCCCTGTAAAGTAACCTTCAGAAAAA---GTTGAGTCAGAT 1920
Qy 275 GlyLeuIlePheAspProAlaGlnIleThrSerArgThrAla----- 288
Db 1921 AAACAAAAATATGATAAATATGTAGCTAAGACGCAACAGCTCTCAAAATAAACAATTAGAA 1980
Qy 289 -----ArgGlyValAlaValProHisGlyAsnHis 298
Db 1981 CAAGAAAAACAAAATGATAGTTGTCAAAACAGGAACTGCATCTAAATCATCTGAT--- 2037
Qy 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArg----- 316
Db 2038 -----GAAATGTATCATCAACAACAATAATCAATGCCTTAATTATTTCA 2079
Qy 317 -----IleIleProLeu 320
Db 2080 AAGTTTCATTAATACTATCAAAATTCAAAATATTATGCTTCACAAATGTTTGAAGAAAT 2139
Qy 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340
Db 2140 AGACGTCAACGAGAACCGTAAAGTGTCTCAAAAGCGTCGATTATAAAAGCGTTGCAACAA 2199
Qy 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
Db 2200 AAGCGTCAAGAACATNAAAAACGAGAG-----CAAGATGCAATACAACT 2244
Qy 361 ProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
Db 2245 GCATTGTGAAATGTATGTCTAAACAGCGAAGCGTATGTTGGTGATAGTTCAATTAAT 2304
Qy 377 ----- 377
Db 2305 GATGATAGTGACTTAAACAGATAATAGTACAGATGCTAGTCAGCTTCATACAAATGCGATA 2364
Qy 377 ----- 377
Db 2365 GAGAAATGAAACTGTATCAAAATGATGAAAAATAAACAAGCGTCAATACAAAATGAAGACACT 2424
Qy 378 -----TyrValPheGluGluAsnGlyValSerArgTyr 388
Db 2425 AATGACACTCATGTAGATGAAGTCCATCAATATAGAGGAAGTTAGTTTGAATCAAGTA 2484
Qy 389 IleProAlaLysAspLeuSerAlaGlu-----ThrAlaAlaGlyIleAspSerLysLeu 406
Db 2485 TCGACACAAAAACAATTTGTAGATGATGAAGTTACGGTTTCGAAATGTAACGCTCAAA--- 2541
Qy 407 AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys---LysThrAspLeuProSer 425
Db 2542 ---CATCAATCAGCACTACACATCAAGCTTGAAGTAAATGATGAAGATGAACATAAAAT 2598
Qy 426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445
Db 2599 CATCCGAA-----TTAATTGCTGATTCAGAGAAGATGGA 2634
Qy 446 LeuAspAsnLys-----GlyArgGlnValAsp-----PheGluAla 457
Db 2635 GCAACGATTAAGAGATATTCAGKAAGTCAATCGATGATGAGATTTTATGAATTA 2694
Qy 458 LeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal*** 477
Db 2695 AATGATACA-----GAAGTAGATGAGGATACCTACTTCAAAATATCGAA 2736
Qy 478 AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAla 497
Db 2737 GATAATACCAATAGAAACGCGTCTGAAATGCAAT-----GTAGACGCTCTCTAAAAAG 2787
Qy 498 GlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThr 517

```

2788 CAA---GAGTACGCACTAAGTAACTCAAGTAAATAATATC-----GATAAAACGGTT 2838
518 GluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVal 537
2839 GATAATGAAATTCAATTAGCAGCGCTCATATAAAAGATGACCAAACTTAAGTGTC 2898
538 -----ThrProHisMetThrHisSerHisGlyTyrIlePheAspSerLeuSer 553
2899 AACTCATTTGAAACGAATGATGTAATGATGTAATGATGTAATGTAATGTAATGTAAT 2958
554 GluAlaGluArgAlaAlaGlnAlaTyrAlaGlyGluGlyGlyLeuThrProProSer 573
2959 GAAATGAAAGAAATGACCGGAAATGACGAAATGACGAAATGACGAAATGACGAAATG 3018
574 ThrAspHisGlnAspSerGlyAsnThrGluAlaGlyGlyAlaGluAlaIleTyrAsnArg 593
3019 GAACAA-----AATGTCGAGAGAAACTATTGAAACGTAATCCAAAG 3063
594 ValLysAlaAlaLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrVal 613
3064 AAACAGACTGAAAGGTTTCACTTAAAGTAAAGACCAATTAATGTTGTCATGACGCCA 3123
614 -----GluValLysAsnGlySerLeuIleIleProHisTyr 625
3124 TCTGATAAAAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATTA 3183
626 AspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAlaProLys 645
3184 AAGCGTGACAAAGT-----GluValLysAsnGlySerLeuIleIleProHisTyr 625
646 GlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665
3202 CAAGCTGTGATGAAAGAAATGCTGCGAGT-----CAAGCCACACCATCATCA 3249
666 ArgProHisSerAspAsnGlyPheGlyAsnAla-----SerAsp 678
3250 AGATCTGATGATCAAGAGTCAAAATACAAATGCAATATAAAACAAATATATGACATCAAC 3309
679 HisValGlnArgAsnLys---AsnGlyGlnAlaAspThrAsn----- 691
3310 AATGTTGAGACAACTCACTTATGTTGTCGAGAACACAGAAATGATTATCAAAATGCA 3369
692 -----GlnThrGluLysProSerGluGluLysProGlnThrGluLysProGluGlu 708
3370 CAACATATTCAGACGACAGAACCTTCTGTTGATCACTCAACGGAATATTTGAAGAA 3429
709 GluThrProArgGluGluLysProGlnSerGluLysProGluSerProLysProThr--- 727
3430 AGTCAAGATGATAATCAATTGGAAATGACGAAATGATGATCAATCAACTCTGCTTCTAGTT 3489
728 GluGluProGluGluSerProGluGluSerGluGlu-ProGlnValGluThrGluLysVa 747
3490 TCAGAAGTAAGCACAATACTGAAAGAAAGCAGAAACACACATCCAAACATACTAGT 3549
747 lGluGluLysLeu 751
3550 GCACACAAAGATA 3562

RESULT 27

US-09-036-987A-1/c
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production

NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
SEQUENCE CHARACTERISTICS:
LENGTH: 80161 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1
Alignment Scores:
Pred. No.: 0.00896 Length: 80161
Score: 159.50 Matches: 158
Percent Similarity: 30.55% Conservative: 87
Best Local Similarity: 19.70% Mismatches: 284
Query Match: 3.97% Indels: 275
DB: 3 Gaps: 27
US-09-765-272A-66 (1-763) x US-09-036-987A-1 (1-80161)
Qy 187 HisGlyAspHisTyrHisTyrIleProLysAsnGluLeu-----Ser 200
Db 44150 CACGCTCCCAACGCAACCGCGTCCCGACCAATCCCTCGCCGCGCTCGGCCAAC 44091
Qy 201 AlaSerGluLeuAlaAlaGluAlaTyr----- 210
Db 44090 GCATCAAGTAGCGCTTCGACGCGCATACGCACCTGTCCGCGGCTCCCAACCCCC 44031
Qy 211 -----TyrAsnGlyLysGlnGlySerArgProSerSerSerSer 224
Db 44030 GACACCGAGGAGAACATCAAAACGATCCAGTCCGAGTCCGCGGCGCAACTGCTCCAGA 43971
Qy 225 TyrAsnAlaAsnProAlaGlnProArg---LeuSerGluAsnHisAsnLeuThrValThrP 244
Db 43970 TTC-----GCGCGACACGACCTTCGCGCGCAACACCGCTGCAAAATCCGGCC 43923
Qy 244 roThr-----TyrHisGlnAsnG 250
Db 43922 AAACCATCTCCACTACGTCCGCGCTTCGCGGACCCCGCGCGGTGTACCAACCGAGTC 43863
Qy 250 In----- 250
Db 43862 AGCGACACTCCCAACCGATCCCTGCCAACACTCCAGCCACTGCTCAGTACGACACG 43803
Qy 251 -----GlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu-----S 267
Db 43802 TCGCAGGCGACATCGACACCGGAGCACCAGCTCCAGTCCGCGCGCAATTCCTCCA 43743
Qy 267 erGluArgHisValGluSer---AspGlyLeuIlePheAspProAlaGlnIleThrSerAr 286

43742 GCGCCCGGAGCGTCGCGTCTCTCGACGCGGTGGTCAACACACAGCTGCTCAGCCCGCGCACCG 43688

286 gThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyr----- 304

43682 GCCAACACCGGGCAACATCGCGCCCCCAACCGCCCGGTACCAACCGTGCACCAACACCGTC 43623

305 ----GluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrAr 323

43622 CCGCGAGGACGCGCACACCGACCGCGCACCGACCGACCGACCGACCGGTCTCCAACCGCGCGACG 43563

323 gSerAsnHis-----TrpValProAspSerArgProGluGlnProSerProGlnSe 340

43562 AACACACACCGCGGANAATGGGACATTGATCTCGCCACCATTCGCCCAACACACCGCCGCC 43503

340 rThrProGluProSerProSerProGlnPro----- 350

43502 AACCGGGCACCAATCGGTCTCCAGTCCGATGGCAGGTGCATCAAGCACCCCGAGAAC 43443

351 -----AlaProAsnProGlnProAlaProSer-A 360

43442 TCCGGGTGCTCCAGTCCCGCGCACGACGACGAGCGCCCGACACGAGCGCTGTCTCCGGAGC 43383

360 snProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValP 380

43382 AC-----GGTACATCT 43371

380 heGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaA 400

43370 TCCAACGCGACGCGCGCGGTACCGGCCCAAAAGGCGCTCGACGTCGCGCTTCGCG 43311

400 laGlyIleAspSerLysLeuAlaLysGlnGlnSerLysSerHisLysLeuGlyAlaLysL 420

43310 AAGGC---CCGACGAAGCTGCACAGTCAGCGCGCAACACAGTTGGCAACGACGAGAATCC 43254

420 ysThrApp-----LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuL 438

43253 GGACGCGCGCATTCATCGAGCCCCAACAGCGCAAGCACGCCACCGAGTTGACCGCGCTCC 43194

438 euAlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaL 458

43193 GTCAGCGCATTTGGCCAGCGCTGTGCGACGCTCTCGGGTGGGGTCTCTTCGACAGG 43134

458 eu-----AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuV 476

43133 CCGATCCGAGACGACTCGG-CGCGCGCTCGACCGACGAGTAACCTGTCAAAGCGG----- 43082

476 al***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProA 496

43081 -----TGACCAATCCGTATCGCGAGCACCGGGCACGACCA 43045

496 snAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAla----- 512

43044 CAGCCAGTACCGGAATCTTCGGGCGCGGCACCTCCCGAGAGGATGCCAAGTATTCG 42985

513 -----GlyLysTyrThrThrGluAspGly----- 520

42984 GTACCGGCAGGAAATCGACGAGCAACGGTTTGGGAGTCCGCTTCGCCAGGAGCAACAGC 42925

520 ----- 520

42924 CGGCAATATACGTCTCAACGATGGCGCATGTGTCGTCTGCGCCCCCAAGATCCAGACGCGC 42865

521 -----TyrIlePheAspProArgAspIleThrSerAspGluGlyAspA 535

42864 GACCACGTTGCCAGATCGCTTGGTCAACCGCGGTCCAGAACTGTTTCGCCACCTCGCC 42805

535 laTyrValThrProHisMetThrHisSerHisIlePheLysLysAspSerLeu----- 552

42804 GATGCTGCGGGTGGCTGCTGGCCGCGCATCGGACTCGATCCAGTAGTGTGTCGTG 42745

553 -----SerGluAlaGluArgAlaAlaGlnAlaLafTyrAlaLysGluLysGlyLeuT 570

42744 GAAGCAATACGTTGGCACTCGAGTCTGATGCGCGCGGTACCGCGCAAAAAGCCCTGCCA 42685

Qy	570	hrProProSerThrAspHisGlnAspSerSerGlyAsnThr	-----582
		:::	
Db	42684	GTCCACCT-----CACCAACACGGGTGTGAGATCGCCAGGGCTGTCTATCATCCGAAACG	42631
Qy	583	-----GluAlaLysGlyAlaGluAlaIleTyrA	592
Db	42630	GGTCTCGTCCCGCTCCCTGCGCATCAGTGGGATCGAGAAATCCGCCCAACCCGATCGGA	42571
Qy	592	snArgValLysAlaAlaLysLysValPro-----LeuAspArg	604
Db	42570	CTCGGTGCGACACTCCCGAACCAACAGTGCGCCAGCGTTCCGTCCGGCGCGAGCTCGACGAC	42511
Qy	605	MetProTyrAsnLeuGlnTyr-----ThrValGluValLysAsnGly	618
Db	42510	CGTGCCTATGCCCTCGCTGCCAGCGCCTGGACACCGTCCGGCGAAACGGACGGGTTCACG	42451
Qy	619	-----SerLeuIleIleProHisTyrAspHisTyr	628
Db	42450	CGCTGGCGGCCCAACATACCTTGGTGTGTCATTTCCGGGGCTCTGTCTCGAGCTCACCGGT	42391
Qy	629	-----HisAsnIleLysPheGluThrPheAspGluGlyLeuTyrGluAlaProLysGly	646
Db	42390	CAACGTTCGACACACAGCGGTGTCTGGCTTACCGAACTCCGACAGCTCGGCAGTCTGGGC	42331
Qy	647	TyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg	666
Db	42330	GAACCTCGCCACACATCGGTCTCATCCGTGCGAATGAAAGCATGGGAGACCCGCAAGCG	42271
Qy	667	ProHisSerAspAsnGlyPheGlyAsnAlaSerAsp-----	678
Db	42270	CCGCTCGGAGTCTCGAGCTCGTTCAGCCGCGCAGATGCTGGCCAGACCTCCCGATC	42211
Qy	678	-----	678
Db	42210	GCCCGACACACACCGAACCGGGGCGATTGAGCGCGCGACGCTCACGGACTCCTCCCG	42151
Qy	679	-----HisValGlnArg	682
Db	42150	ACCGGCCAGCAATGCGCCGAACAGGTCTTCACGACGAGCCACCGCAACATCG-CCGCGT	42092
Qy	683	AsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGlu-----	698
Db	42091	CAGAGGGCAGGCTTGATCAACCGGGGCGCGGCCACCAACCGAGCCGATCCCGCA	42032
Qy	699	-----GluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLys	715
Db	42031	ACGACAAACACGCGCAGCGCGGCGCGGCGCA-----ACTCCCGACCGAATGCC	41981
Qy	716	ProGlnSerGlu-----LysProGluSerProLysProThrGluGluProGluGluSerProGln	735
Db	41980	CCATCACCACATCCGGCCGAAACGCCCAACA-----ACCAGAGCCCAAGAGCGCGGCTT	41924
Qy	735	uGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArg	752
Db	41923	GCAGCGCAAGACACCGCTCGCCACACAGCTCTGATCCAGCAGCTGCGC	41872

RESULT 28

REFSOL 28
US-09-370-700-1/c

: Sequence 1, Application

Sequence 1, Application 08/053763
Patent No. 6274350

; GENERAL INFORMATION:

; APPLICANT: Baltz, Richard H

APPLICANT: Broughton, Mary C

APPLICANT: Crawford, Kathryn P

APPLICANT: Madduri, Krishnamurthy

APPLICANT: Treadway, Patti J

APPLICANT: Turner, Jan R

APPLICANT: Waldron, Clive

; TITLE OF INVENTION: Biosynthetic

; FILE REFERENCE: 50489 DIV1

; CURRENT APPLICATION NUMBER: US/09

; CURRENT FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/36987

NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 80161

TYPE: DNA

ORGANISM: Saccharopolyspora spinosa

US-09-370-700-1

Alignment Scores:

SeqScore: 0.00896 Length: 80161

Percent Similarity: 30.55% Matches: 158

Best Local Similarity: 19.70% Mismatches: 87

Query Match: 3.97% Indels: 275

DB: 3 Gaps: 27

US-09-765-272A-66 (1-763) x US-09-370-700-1 (1-80161)

201 AlaSerGluLeuAlaAlaGluAlaTyr-----210
 44090 GCATCCAGTAGCGTTCGACGGCGATACGACCTGTCCGGCGCTCCACACCCCC 44031
 211 -----TTPAsnGlyLysGlnGlySerArgProSerSerSerSerSer 224
 44030 GACACCGAGGAGAACATCACAACGCGATCCAGTCCGAGTCCGCGCAGCAACTCGTCAGA 43971
 225 TyrAsnAlaAsnProAlaGlnProArg--LeuSerGluAsnHisAsnLeuThrValThrP 244
 43970 TTC-----GCCGACACCGCACCTTCCGCGACACACCGCTCGCAAAATCCGCC 43923
 244 roThr-----TyrHisGlnAsnG 250
 43922 AAACCCATCTCCACTACGTCCGCGCGCTCGCGACCCCGCGGCGGTATCCACCGCAGTC 43863
 250 In-----250
 43862 AGCGGACACTCCCGACCGATCCCTGCGCAACACTCCAGCCACTCGCTCAGCATCAGCCACG 43803
 251 -----GlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysPProLeu-----S 267
 43802 TCGAGGGGACACATCGACACCGGAGCAGCCCGCGCTCCAGTCCGCGCGCAATTCGCCA 43743
 267 exGluArgHisValGluSer--AspGlyLeuIlePheAspProAlaGlnIleThrSerAr 286
 43742 GCGCGGAGCGTCCGCTCTCGACGGTGGTCAACACACGAGTGTCTCAGCCCGGACCG 43683
 286 gThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyr----- 304
 43682 GCCAACACCGCGCAACATGCGGCGCCCAACCGCCCGTACCAACCGCTCAACACACCGCTC 43623
 305 -----GluGlnMetSerGluLeuLysArgIleAlaArgIleIleProLeuArgTyrAr 323
 43622 CCGGAGGAGCGCCACCGACCGGACCGCGACCGCGACCGCGCGGTTCCAAACCGCGCAGC 43563
 323 gSerAsnHis-----TTPValProAspSerArgProGluProSerProGlnSe 340
 43562 AACACACACCGCGGATGGGACTTGTCTCGCCACCATCCGCGCAACACACCGCGC 43503
 340 rThrProGluProSerProSerProGlnPro-----350
 43502 AACCGGCGCCCAATCGGTCTGCCAGTCCGATGCGAGGTGATCAAGCCACCCAGAAC 43443
 351 -----AlaProAsnProGlnProAlaProSer-A 360
 43442 TCGCGGTGTCTCCAGTCCCGGACACCGCAGCGCGCGCGCGCGCGTCTCGCGAGAC 43383
 360 snProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValP 380

43382 AC-----CGGTACATCT 43371
 380 heGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaA 400
 43370 TCAACGCGACCGCGCGCGCTCGCCGCAAAAGCGCTCGAGTCCGCGCTCCGC 43311
 400 laGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysL 420
 43310 AAGGC---CCGCAAGAGCTGCACAGTCAGCGCGAAACAGTTGTCAGCAGGAGAAATCCC 43254
 420 yThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuL 438
 43253 GGAGCAGCGCGATTCATCGAGCCCCAACAGCGAAAGACCGCACCGAGTTGACCGCATCC 43194
 438 euAlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaL 458
 43193 GTCAGCGCATTCGCGCGCTGTCGCGCTCTCTCGCGCTCTCTCTCGGACAGG 43134
 458 eu-----AspAsnLeuLeuGluArgLeuLysAspVal**SerAspLysValLysLeuV 476
 43133 CCGATCCGAGACCTCGG-CGCGCTCCGACCGAGTGAATCTCTCAAGCGC-----43082
 476 al***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProA 496
 43081 -----TGACCAATCCGTCATCCGACGCGCGCGCGCACGACCA 43045
 496 snAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAla-----512
 43044 CAGCCAGTACCGGAAATCTTCGCGCGCGCGCACCTCCGAGAGGAATCCCACTGATTCG 42985
 513 -----GlyLysTyrThrThrGluAspGly-----520
 42984 GTACCGCGAGATCGAGCGGACGAGTTCGCGAGTCCGCTTCGCCAGGAGCAACCGC 42925
 520 -----42925
 42924 CGGCAATACATCGTCAACGATGCGATGTGTCTGTCGCCCCCAAGATCCAGAGCGCG 42865
 521 -----TyrIlePheAspProArgAspIleThrSerAspGluGlyAspA 535
 42864 GACCAACGTTCCGATCGCTTGGTCAACCGCGCTCAGAACTGTCGCCACCTCGCGC 42805
 535 laTyrValThrProHisMetThrHisSerHisTyrIleLysLysAspSerLeu-----552
 42804 GATGTCGCGGCTCGCTGCTCGCGCGCTCGACTCGATCCAGTAGTGTGTGCTGTG 42745
 553 -----SerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuT 570
 42744 GAAGGCATACGTTGGCAACTCGAGTCTAGCGCGGTACCGCGGCAAAACCGCTGCCA 42685
 570 hrProProSerThrAspHisGlnAspSerGlyAsnThr-----582
 42684 GTCCACCT-----CACACACCGGTTGGAGATGCCCGAGGCTGTCTATCACCAGAG 42631
 583 -----GluAlaLysGlyAlaGluAlaIleTyrA 592
 42630 GGTCTCGTCCGCTCCCTGCGCATCAGTGGGATCGACGAATCCGCCAACCCGATCGGA 42571
 592 snArgValLysAlaAlaLysLysValPro-----LeuAspArg 604
 42570 CTGGTCCGACACTCCCGAACACAGTCCGCGCGCTCCGTCGCGCGCGAGTCCGACGAC 42511
 605 MetProTyrAsnLeuGlnTyr-----ThrValGluValLysAsnGly-----618
 42510 CGTGCCTATGCTTCGCTCCGAGCGCTCGACACCGTTCGCGGAAACCGAGCGGTTCACG 42451
 619 -----SerLeuIleIleProHisTyrAspHisTyr-----628
 42450 CGCTGGCGGACCCCAATACCTGTGTGCTCTATTCGCGGCTCTGTGAGCTACCCGT 42391
 629 -----HisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAlaProLysGly 646

Db 42390 CAACGTCGACACAAAGCGGTGCTGGTTCACGAACCTCCGACGACTCGCGCATCTGGGC 42331
2Y TyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg 666
Db 42330 GAATCGCCCAACATCGGTTCATCGGTHGCGAATGAAGAAGCATGGAGACCCGCAAGCG 42271
2Y ProHisSerAspAsnGlyPheGlyAsnAlaSerAsp- 678
Db 42270 CCGCGTCGCGACTCGGAGCTCGGTTCAGCCGCGCGAGCATGCTGGCCAGCACCTCCCGATC 42211
2Y 678 ----- 678
Db 42210 GCCCGACAAACACCAACCGAACCAGCGGCGCATTTAGCGCGCGGAGCGCTCACGAGCTCTCTCCCG 42151
2Y 679 -----HisValGlnArg 682
Db 42150 ACCGCGCAGCAATGCGCGCAACGAAGTCTTCAACGACGACGACCGCCAACTCG-CCGCGT 42092
2Y 683 AsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGlu- 698
Db 42091 CAGAGGCGAGGCTTGATCAACCGGCGCGCGCGCCACCAACGAGCGCATCCCGCA 42032
2Y 699 -----GluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLys 715
Db 42031 ACGACAAACAGCGCGCAACCGCGCGGCA-----ACTCCCGACCGAATGCC 41981
2Y 716 ProGlnSerGlu-LysProGluSerProLysProThrGluGluGluProGluGluSerProGln 735
Db 41980 CCATCACCACATCCGCGCGCAACCGCCCAAG-----ACCCAGCAGCCCAAGAGCGCGGCTT 41924
2Y 735 uGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArg 752
Db 41923 GCAGCGCGAAGACAGACCGGACTGCGCGCCACAGGCTGTGATCCAGCAGCTGGCG 41872

RESULT 29

US-09-603-207-1/c

Sequence 1, Application US/09603207B

Patent No. 6521406

GENERAL INFORMATION:

APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/603,207B
CURRENT FILING DATE: 2000-06-23
EARLIER APPLICATION NUMBER: 09/370,700
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
TYPE: DNA
ORGANISM: Saccharopolyspora spinosa

Alignment Scores:

Alignment No.:	0.00896	Length:	80161
Score:	159.50	Matches:	158
Percent Similarity:	30.55%	Conservative:	87
Best Local Similarity:	19.70%	Mismatches:	284
Query Match:	3.97%	Indels:	275
DB:	4	Gaps:	27

US-09-765-272A-66 (1-763) x US-09-603-207-1 (1-80161)

2Y 187 HisGlyAspHisTyrHisTyrLeuProLysAsnGluLeu-----Ser 200
Db 44150 CACGGTCCCAACGCAACCGCGGTCCCGACCAATCCCTCGCCGCGAGCTGCTCGGCCAAC 44091

QY 201 AlaserGluLeuAlaAlaAlaGluAlaTyr----- 210
Db 44090 GCATCCAAGTAGGCGTTTCGACGCGCATACGACCCCTGTCGCGCGCTCCCCACACCCCC 44031
QY 211 -----TrpAsnGlyLysGlnGlySerArgProSerSerSerSerSer 224
Db 44030 GACACCGAGGAGAACATCAACAACGCATCCAGCTCCGAGTCGGCCAGCAACTCGTCCAGA 43971
QY 225 TyrAsnAlaAsnProAlaGlnProArg--LeuSerGluAsnHisAsnLeuThrValThrP 244
Db 43970 TTC-----GCCGACACCGACCTTCGCCGACAACACCGCTGCAAAATCCGCC 43923
QY 244 rothr-----TyrHisGlnAsnG 250
Db 43922 AAACCCATCTCCATACGTCCCGCGCTCGCGACCCCGCGCGCGTGTACCAACCGCAGTC 43863
QY 250 ln----- 250
Db 43862 AGCGGACATCCCCACCGATCCCTGCCAAACATCCAGCCATGCTGCTCAGATCAGCCACG 43803
QY 251 -----GlyGluAsnLysSerSerLeuLeuArgGluLeuTyrAlaLysProLeu-----S 267
Db 43802 TCGCAGGCGCAATCGACACCCGAGCACCAGCGCTCCAGCTCCGCCCGCAATTCGCCA 43743
QY 267 erGluArgHisValGluSer--AspGlyLeuIlePheAspProAlaGlnIleThrSerAr 286
Db 43742 CGCGCCGAGCGTCCGTCTCTCGAGCGTGTCAACACCGTGTTCAGCCCGCGCACCG 43683
QY 286 gThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyr----- 304
Db 43682 GCCAACACCGGCGCAACATGCGCGCCCAACCGCGGTACCAACCGCTACCAACACCGTC 43623
QY 305 ----GluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrAr 323
Db 43622 CCCGAGGACGCCACACCGACCCCGCACCGCAGCGCTTCCAAACCGCGCACG 43563
QY 323 gSerAsnHis-----TrpValProAspSerArgProGluGlnProSerProGlnSe 340
Db 43562 AACACACCCACCGCGAATGGCGACTTGATCTCGCCACCATCCGCCAACACACCGCGC 43503
QY 340 rThrProGluProSerProSerProGlnPro----- 350
Db 43502 AACCGGCGACCCCATCGGTCTGCCAGTCCGATGCGAGCTCGATCAGCCACCCAGAAC 43443
QY 351 -----AlaProAsnProGlnProAlaProSer-A 360
Db 43442 TCCGGTGTCTCCAGTCCCGCGACACGACGAGCCCGCCAGACCGAGCGCTCTCCGAGAC 43383
QY 360 snProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValP 380
Db 43382 AC-----GGGTACATCT 43371
QY 380 heGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaA 400
Db 43370 TCCAAACGCGAGCGCGCGGTCTACCGCCCAAAAGGCGCTCGACGTCGCGCTCCGC 43311
QY 400 laGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysL 420
Db 43310 AAGGC---CCGCAAGAAGCTGCACAGTCAGCGCGAAACCCAGTTGGCAAGCAGGAGATCC 43254
QY 420 ysThrAsp-----LeuProSerSerSerArgGluPheTyrAsnLysAlaTyrAspLeuL 438
Db 43253 GGAGCAGCGGATTCATCGAGCCCAACAGCAGCAAGAGAGCGCCACCGAGGTGACCGGCATCC 43194
QY 438 euAlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaL 458
Db 43193 GTCAGCGCATTTGCCGCGCTGTGGCAGCTCTCGCGGTGCGGGTCTCTTCGACAGCAG 43134
QY 458 eu-----AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuV 476
Db 43133 CCGATCCGAGCAGCCTCGG-CGCGCGCTCCGACCACTGACTCGTCAAGCCG----- 43082

476 aI**AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProA 496
43081 -----TGACCAATCCGTATCGCAGCACCGGCGCACGACAA 43045
496 snAlaGlnIleThrTyrThrAspAspGluLeuGlnValAlaLysLeuAla- 512
43044 CAGCCAGGTACCGGAATCTTCGGGGCGGCACCTCCCGAGAGGAATGCCAACTGATTGC 42985
513 -----GlyLysTyrThrThrGluAspGly----- 520
42984 GTACCGCGAGGATCGACGAGCGAAGCGTTTCGGAGTCCGCTTCGCCAGGAGCAACGC 42925
520 ----- 520
42924 CGGCAATACATCGCTCAACATCGCATGTGTCTGCTGCCGCCCAAGATCCAGAGCGGC 42865
521 -----TyrIlePheAspProArgAspIleThrSerAspGluGlyAspA 535
42864 GACCAAGTTGCCAGATCGCTTGGTCAACCGCGGTCCAGAACTGTTCCGCCACCTCGCC 42805
535 laTyrValThrProHisMetThrHisSerHisIleThrLysLysAspSerLeu----- 552
42804 GATGTCTCGCGGTGCGGTGTCGCCGCGCACTCGCACTCGATCGATGCTGTGTGTTG 42745
553 -----SerGluAlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuT 570
42744 GAAGGCATACGTTGGCACTCGAGCTGCTAGCGCGGTACCGGAAACGCGCTGCCA 42685
570 hrProProSerThrAspHisGlnAspSerGlyAsnThr----- 582
42684 GTCCACT-----CACCAACCGGTGTGAGATGCCAGGCTGTCTATCCGGAACG 42631
583 -----GluAlaLysGlyAlaGluAlaIleTyrA 592
42630 GGTCTCGTCCGCTCCCTGCGCATCAGTGGATCGACGAATCCGCCAACCCGATCGGA 42571
592 snArgValLysAlaAlaLysLysValPro-----LeuAspArg 604
42570 CTCGGTCGCACACTCCCGAACACGATGCGCCGAGCTTCGTCGGGCGCGAGCTCAGAC 42511
605 MetProTyrAsnLeuGlnTyr-----ThrValGluValLysAsnGly----- 618
42510 CGTGCCTATGCTCGCTGCGCGCCGCTGGACACCGTGGCGAAACGCGGTTCAAG 42451
619 -----SerLeuIleIleProHisTyrAspHisTyr----- 628
42450 CGCTGGCGCACCAATACCTCGTGTGTCTATTTCCGGCGCTGTGTGAGCTACCCGT 42391
629 -----HisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLysGly 646
42390 CAACGTCGACACACGCGGTGTGTGGCTTACCGAACTCCGCACTCGCGAGCTCGCGCATCTGGGC 42331
647 TyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg 666
42330 GAACTCGCCCAACATCGGTCTCATCCGTGCGAATGAAGCATCGGACCGCGAAGCG 42271
667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAsp----- 678
42270 CGCGTCCGGACTCGAGTCTCGTCAAGCGCGCGACGATGCTGGCCAGCACCTCCCGATC 42211
678 ----- 678
42210 GCCCGACACACCGAACCGGGGCGATGTGAGCGGCGAGCTCAGGACTCTCCCG 42151
679 -----HisValGlnArg 682
42150 ACCGGCAGCATGGCCGAACAGGTCTTCACCAGCAGCCCGCCCAACATCG-CGCCGT 42092
683 AsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGlu----- 698
42091 CAGAGGCGAGGCTTGCATCAACCGCGCGCGCGCCACCAACCGAGCGCGATCCCGCA 42032
699 -----GluLysProGlnThrGluLysProGluGluThrProArgGluGluLys 715

42031 ACACAACACGCCAGCGCAACACGCGGCCA-----ACTCCCCGACCGAATGCC 41981
716 ProGlnSerGlu-LysProGluSerProLysProThrGluGluProGluGluSerProG 735
41980 CCATCACACATCGCGCGAGCCGCCACGA---ACCAGCAGCCCAAGAGCGCGGCTT 41924
735 uGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArg 752
41923 GCAGCGGACACAGACCGCGACTGCGGCCACAGCGTCTGATCCAGAGCTGCGC 41872
RESULT 30
US-09-286-981B-29
; Sequence 29, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenmann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-29
Alignment Scores:
Pred. No.: 1,67e-05 Length: 1278
Score: 158.00 Matches: 128
Percent Similarity: 36.03% Conservative: 77
Best Local Similarity: 22.50% Mismatches: 183
Query Match: 3.93% Indels: 181
DB: 4 Gaps: 29
US-09-765-272A-66 (1-763) x US-09-286-981B-29 (1-1278)
QY 216 GlySerArgProSerSer-SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSe 235
DB 12 GGTAACTACCCAGTAGCCACTTCTTCTAATAAGGCAATAAAGTCAG----- 60
QY 235 rGluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSe 255
DB 61 -----ACAGAACATATGAAAGCTGCTAAACCAAGTCGA 92
QY 255 rSerLeuLeuArgGluLeuTyrAlaLysProLeuSer-----GluArgHisValG 272
DB 93 TGAATATATAAA-----AAAAAGCTCCAATTAGATAGAGAAACATACC 140
QY 272 uSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAl 292
DB 141 AAATGCGGCTTACTCACAAGTTGGCGTAAATAAAGCGGAGTATTTGTCATGGATTAG 200
QY 292 aValProHisGlyAsnHisTyrHisPheIleProTyr-GluGlnMetSerGluLeuGly 312
DB 201 TGTTCAAAAAGAGTCGGAAGCTGAGTGTGCCG-----TCAGAAATAAAGC 248
QY 312 sArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerAr 332
DB 249 AAGATT-----GACGCAGC 263
QY 332 gProGluGlnProSerProGlnSerThrPro---GluProSerProSerProGlnProAl 351
DB 264 TTTTGACGAGTTTAAAAAAGATACATTACCACAGAACAGGAAAAAGGTAGCAGAACG 323

[illegible]

```

QY      690 rAenGlnThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGlu-- 709
Db      1083 AGAAGCTAAACGAAGAGCAGCAGAGATAAAGTTAAAGAAAAAACCCAGCTGAACAAC 1142
QY      710 -----ThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725
Db      1143 ACNACCAGCGCGGCTCTCAACAGAAAAACCA---ACTGAGAGCTTGAGATCCAGC 1199
QY      725 sProThrGluGluProGlu---GluSerProGluGluSerGluGluProGlnValGlu-- 743
Db      1200 TCCAGCTCCAGCTCCAAACACCAGAGAATCCA-----GCTGAAAAACCAAAAGCAGAAAA 1253
QY      744 -----ThrGluLysValGluGlu 749
Db      1254 GCCAGCTGATCAACACAGCTGAAGAA 1278

RESULT 31
US-09-286-981B-31
; Sequence 31, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-31

Alignment Scores:
Pred. No.: 1.85e-05 Length: 1272
Score: 157.50 Matches: 126
Percent Similarity: 35.89% Conservative: 75
Best Local Similarity: 22.50% Mismatches: 181
Query Match: 3.92% Indels: 178
DB: 4 Gaps: 28

US-09-765-272A-66 (1-763) x US-09-286-981B-31 (1-1272)
QY      216 GlySerArgProSerSer-SerSerSerTyrAsnAlaAanProAlaGlnProArgLeuSe 235
Db      12 GGTATACCCAGTAGCAGCTCTCTTAATGGGCATATAAAGTCAG----- 60
QY      235 rGluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyAsnIleSe 255
Db      61 -----ACAGACATATGAAGCTGTCTAAACAAGTCGA 92
QY      255 rSerLeuLeuArgGluLeuTyrAlaLysProLeuSer-----GluArgHisValGln 272
Db      93 TGAATATATAAAA-----AAAAAGCTCCAATTAGATAGAAGAAAAACATACCCA 140
QY      272 uSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAl 292
Db      141 AAAATGCGCGCTTACTCACAAGTGGCGCGTAATTAAACGAGGATVTTGCATGGATTAG 200
QY      292 aValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLys 312
Db      201 TGTTCAAAAAGAGAGCTCGAGCTGAGTGTGCG-----TCAGATAATAAAGC 248
QY      312 sArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSer 332
Db      ::::

```

Db 249 AAGTTA-----GACCCAGC 263
Qy 332 sProGluGlnProSerProGlnSerThrPro---GluProSerProSerProGlnProAl 351
Db 264 TTTTGGCAGGTTTAAAAAGATACATTACCAACAGAAAGGTTAAAAAGGTAGCAGAAGC 323
Qy 351 aProAsnProGlnProAlaProSerAsnProIleAspGluLeuValVallyGluAlaVa 371
Db 324 T-----GAGAAAGAGGTTTGAAGAAGCTAA 347
Qy 371 lArglyValGlyAspGlyTyrValPheGluGluAenGlyValSerArgTyrIleProAl 391
Db 348 GAATAAGCCGAGGATCAA-----AAAGAAAAGATCTCCGTAACATC---CCAAC 395
Qy 391 a-----LysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLy 408
Db 396 CAATACCTTACAAAACGCTTGAACCTTGACATCTGAGTCC---GATGTGGAAGTTAAAA 452
Qy 408 sGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAspAr 428
Db 453 AGCGAGCTTGAACCTAGTAAAGAGAGAGCTAAGGAA-----TCTCGAGACGA 500
Qy 428 gGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAs 448
Db 501 GAAAAAATTATCAAGCA-----AAAGCGAAAGTTGAGAA 536
Qy 448 nLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVa 468
Db 537 TAAAAAGCTGAGGCT-----ACAGGTTTAAAAACAT 569
Qy 468 l***SerAspLysVallyLeuVal***AspIleLeuAlaPheLeuAlaProIleArgHi 488
Db 570 CAAGACAGATCGTGAATA-----GCAGAAGAGAGCTAAACAGAGAGC 611
Qy 488 sProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVa 508
Db 612 AGATGCTAAGTTGACGAGGAACTAATGTACGACTTCAGACGAAAGATAAATCAAGGCG 671
Qy 508 lAlaLys-----LeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProAr 526
Db 672 GCGAAAAAGAGAGTTCTTGAGAGCTAGCAACCTGAT----- 711
Qy 526 gAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTr 546
Db 711 ----- 711
Qy 546 pIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGl 566
Db 712 -----AAAAAA-----GAAAAATGATCGGAAGTCTTCAGATTTCTAGCGTAGGTGA 755
Qy 566 uLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGl 586
Db 756 AGAACTCTTACAGCCCATCCTCG-----AAACCGAATAAAAGGT 797
Qy 586 yAlaGluAlaIleTyrAsnArgVallyAlaAlaLysLysValProLeuAspArg----- 604
Db 798 AGCAGAAGCT--GAGAGAAGGTTGAAGAGCTAAGAAAAAGCCGAGGATCAAAAGA 854
Qy 605 -----MetProTyrAsnLeuGlnTyrThr----- 612
Db 855 AGAAGATCGTGTACTACCAACCAATACTTACAAAACGCTTGAACCTGAAATTCCTGA 914
Qy 613 -----ValGluVallyAsnGlySerLeuIleIleProHisTyrAspHisThrHisAs 630
Db 915 GTCCGATGTGGAAGTTAAAAAGCGAGCTT----- 945
Qy 630 nIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGl 650
Db 946 -----GAAGTAGTAAAGAGGAGGAGCTAGGATCTCGAAACGA 983
Qy 650 uAspLeuLeuAlaThrVallyTyrTyrValGluHisProAsnGluArgProHisSerAs 670
Db 984 GGAATAAATTAAGCAAGTAAAGCAAGAAAGTTGAG-----AGTAA 1022

Qy 670 pAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspTh 690
Db 1023 AAAAGCTGAGCTACAGGCTAGAAAAATCAACACAGATCGTAAAAAGCAGAGAAGA 1082
Qy 690 xAsnGlnThrGluLysProSerGluGluLysProGlnThrThrGluLysProGluGluGlu 709
Db 1083 AGAAGCTAAAGCAAGCAGCAGCAAGAAAGATAAAGTTAAAGAAAAAACAGCTGAACAACC 1142
Qy 710 -----ThrProArgGluGluLysProGlnSerGluLysProGluSerProLy 725
Db 1143 ACAACAGCCGCGCTCTCTCAACAGAAAAACCA---ACTGAAGAGCTTGAGATCCAGC 1199
Qy 725 sProThrGluGluProGlu---GluSerProGluGluSerGluGluProGlnValGlu 743
Db 1200 TCCAGCTCCAGCTCCAAAAACAGAGATCCA-----GCTGAAAAAACCAAAAGCAGAA 1251

RESULT 32
US-09-286-981B-30
; Sequence 30, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae

US-09-286-981B-30

Alignment Scores:
Pred. No.: 2,32e-05 Length: 1276
Score: 156.50 Matches: 126
Percent Similarity: 35.89% Conservative: 75
Best local Similarity: 22.50% Mismatches: 181
Query Match: 3.89% Indels: 178
DB: 4 Gaps: 28

US-09-765-272A-66 (1-763) x US-09-286-981B-30 (1-1276)

Qy 216 GlySerArgProSerSer-SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSe 235
Db 12 GGTAACTACCAAGTAGCCACTTCTCTTAATAGCAATAATAAAGTCAG----- 60
Qy 235 rGluAsnHisAsnLeuThrValThrProTyrTyrHisGlnAsnGlnGlyGluAsnIleSe 255
Db 61 -----ACAGAACATATGAAAGCTGCTAAACCAAGTCGA 92
Qy 255 rSerLeuLeuArgGluLeuTyrAlaLysProLeuSer-----GluArgHisValGl 272
Db 93 TGAATATATAAAA-----AAAAAGCTTCAATTAGATAGAGAAACATACCCA 140
Qy 272 uSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAl 292
Db 141 AATATCGGCTTACTCACAAGTTGGCGTAATATAAACCGAGTATTTGCGATGATTAAG 200
Qy 292 aValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGly 312
Db 201 TGTTCAAAAAGAGAGTCGGAAGCTGAGTTGCCG-----TCAGAAATAAAGC 248
Qy 312 sArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerAr 332

233 gLueSerGluAsnHisAsnLeu-----ThrValThrProThrTyrHisGlnAsnGlnG1 251
747 ACAAGGAGAACCACTAAATAAATCGATTTCAGAACGAGTAAGCGAAGAGGTGCGA 806
251 yGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisVa 271
807 GGAATATGTAATAAATAATAGTGGGTGAGAGCTATGCAAAAATCAATAAAAGCGACATAC 866
271 lGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyVa 291
867 AATTACTGTAGCTCTAGTTAACGAGTTCAACCACTTAAG----- 906
291 lAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuG1 311
907 -----AACGAGTATTGAAATAAATAGTTGAATCAACCTCAGAAAGCCA 950
311 uLysArgIleAlaArgIleIleProLeuArgTyrArgSer-----As 325
951 A-----CTACAGATACTGATGATGGAGAGTCGATCAAAAGTAGATGAAGCTGTGTC 1001
325 nHisTrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSe 345
1002 TAAGTTTGAAGAGCACTCATCTTCTGTCAGTTCAGACTCTTCCACT---AAACCGGA 1058
345 rProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGlu-- 364
1059 AGCTTCAGATACAGCGAAGCCAAACAGCCAGACAGCAAGAGAAAGGTAGCAGAGC 1118
365 ---LysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluLys 383
1119 TAAGAAAGAGGTGAAGAGCTGAGAAAGCCAGCAAGTCAAA-----AAAGAGA 1169
383 nGlyValSerArgTyr-----IleProAlaLysAspLeuSerAlaGluThrAla--Al 400
1170 AGATCGTCGTAACTACCAACCATTAATTACAAACGCTTGAATTTGAGTCT 1229
400 aGlyIleAspSerLysLeuAlaLysGlnGlnSerLysLeuGlyAlaLysLy 420
1230 CGATGTGGAAGTTAATAAAGCGGAGCTTGA-----CTAGTAAAGTGAA 1274
420 sThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaAr 440
1275 AGCTAACGAACTCGACAGCAGCAAAATAAT----- 1305
440 gIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAs 460
1306 -----AAGCAAGCAAGAGCGGAGTTCGAGAGTAAACAGC 1340
460 nLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLe 480
1341 TGAGGCTACAGGTAAATAAATAATCAAGACAGATCGTGAAGAACAGCAAGAGAA----- 1395
480 uAlaPheLeuAlaProIleArg---HisProGluArgLeuGlyLysProAsnAlaGlnI 499
1396 -----GCTAAACAGAGACAGATGCTAAAGAGCAAGTAAACCAAGGGGCGG-- 1443
499 eThrTyrThrAspGluIleGlnValAlaLys-----LeuAlaGlyLysTyrThrTh 517
1444 -----GCAAACGAGGAGTCTCTCGAGAGCTAGCAAC 1475
517 rGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVa 537
1476 ACCTGAT----- 1482
537 lThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluAr 557
1483 -----AAAAA-----GAAATATGATGC 1499
557 gAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisG1 577
1500 GAAGTCTTCAGATTCTAGCTAGGTAGAGAAACTCTTCCAAGGCCATCCCTG----- 1551

577 nAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAl 597
1552 -----AAACCAAGAAAAAGGTAGCAGAAGCT---GAGAAGAAAGGTGGAAGAGC 1598
597 aLysLysValProLeuAspArg-----MetProTyrAsnLeuG1 610
1599 TAAGAAAAAAGCGAGGATCAAAAAGAAAGAGATCGCGTAACCTACCAACCAATACTTA 1658
610 nTyrThr-----ValGluValLysAsnGlySerLeuI1 621
1659 CAAAACGCTTGAACCTCAAAATTGCTCAGTCCGATGTGGAAGTTAAAAAACGGAGCTT-- 1716
621 eIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTy 641
1717 -----GAACCTAGTAAA 1727
641 rGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValG1 661
1728 AGAGGAAGCTAAAGGAACCTCGAAACGAGGAAAGTAAAGCAAGCAAAAGCGGAAGTTGA 1787
661 uHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValG1 681
1788 G-----AGTAAAGAGCTGAGGCTACAGGTTAGAAAAATCAA 1826
681 nArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysPr 701
1827 GACAGATCGTAAA---AAAGCAGAGAAAGAGCTAAACGAAAGAGCAGCAGAGAGATAA 1883
701 oGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLysPr 721
1884 AGTTAAAGAAAAACAGCTGTAACCAACCAAGCGCGCGCTCCAAAAGCAGAAAAACC 1943
721 o---GluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPr 740
1944 AGCTCCAGCTCAAAACCA---GAGAATCCAGCTGAACCAACCA---AAAGCAGAAAAACC 1997
740 oGlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAsp 756
1998 A---GCTGATCAACAGCTGAGAGAGACTATGCTCGTAGATCAGAGAA 2043

RESULT 34

US-08-961-527-83/c
; Sequence 83, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

Db	486894	CAA---ACTCTC	CAAGAGCTAAATTTGATTCC	ACAGTCGAAATTTCCCAAGAAATCATCG	486893
Qy	299	TyrHisPheIleProTyrGluGlnMetSerGlu	-----LeuGluLysArgIleAla	315	
Db	486897	CAAGCTGAATTTTCATGCTGAA	CAAAATCTCTGATGAAATTAATTTAG	AAGAAAAAACTGAA	486778
Qy	316	ArgIleIleProLeuArgTyrArgSerAsnHis	TrpValProAspSerArgProGluGln	335	
Db	486777	GCAGTTTTT-----GATCAT	CAACAATTAGAAAAATCAGTCTGA	GAAGAA	486736
Qy	336	-----ProSerProGlnSerThrProGluProSerPro	-----SerProGlnPro	350	
Db	486735	ACAGTTGTATACACTACTAGTACTCTCTTTG	AACAGAAACAATTTGAAACCCAAATTA	486676	
Qy	351	AlaProAsnProGlnProAlaProSerAsnPro	-----IleAspGlu	366	
Db	486675	GAACTAGCTCTGAAGATCAACCATCTGAA	CCTGCATTTAGATCAAAATCATTCC	GAAT	486616
Qy	367	ValLysGluAlaValArgLysValGlyAspGly	TyrValPheGlu	381	
Db	486615	GTTACTGCTGAAGTTGAACAAATCTTTG	ATGGAATAAGTTAGAAAGTTTAAACTAG	AG	486556
Qy	382	-----GluAsnGlyValSerArgTyrIleProAla	LysAspLeuSerAlaGlu	397	
Db	486555	GAACTAACTTTGTGATAATGTTGAAATTAAT	GAAGTTCAACCTTAAAGAACAGAACTG	AA	486496
Qy	398	ThrAlaAlaGlyIleAspSerLysLeuAlaLys	GlnGluSerLeuSerHisLysLeuGly	417	
Db	486495	ATTACTTTTGTAT-----GAAACT	TAAGAAGTTCCACAGGAACCTCTCTG	AA	486448
Qy	418	AlaLysLysThrAspLeuProSerSerAspArg	GluPheTyrAsnLysAla	436	
Db	486447	-----CCATTATCAACAGAGAATTTAA	AAATCTGAAGCTACTTTTGAT	486406	
Qy	437	LeuLeuAlaArgIleHisGlnAspLeuLeuAsp	LeuLysGlyArgGlnValAspPheGlu	456	
Db	486405	AATGTATCTGAAGCTGAATCAGAGCAGT	TTTGAAAAA-----CCTCAATTAGAA	CACAG	486349
Qy	457	AlaLeuAspAsnLeuLeuGluArgLeuLys	AspVal***SerAspLysValLysLeuVal	476	
Db	486348	ACTGAAAAATATTGAAAGAGGAACCTTA	ATCTGAACCAAGTTGAT-----CAATTA	ATA	486295
Qy	477	***AspIleLeuAlaPheLeuAlaProIle	ArgHisProGluArgLeuGlyLysProAsn	496	
Db	486294	ACTGAA-----GCTAGCTTTGATCTGT	TAAACATGAAGCAGTTTTTGATA	AAAAATCA	486241
Qy	497	AlaGlnIleThrTyrThrAspAspGluIle	GlnVal-----AlaLysLeuAlaGly	513	
Db	486240	ACTCAA---ACTGAGGGATTGGAAGAAC	CTCAAGTTCTAGTGAAGCTGAAGTAG	TTTGAT	486184
Qy	514	LysTyrThrThr-----GluaspGlyTyr	IlePheAspProArgAspIleThr	529	
Db	486183	CAACAACACTACTGATATCTGTGGAGAAC	CTGAACAGTTTTTTTGAT-----	486139	
Qy	530	SerAspGluGlyAspAlaTyrValThrPro	HisMetThrHisSerHisTrpIleLysLys	549	
Db	486138	-----GTTCAACCTGAAAAAACTACT	GAA-----GTTAAGTT	486106	
Qy	550	AspSerLeuSerGluAlaGluArgAlaAla	GlnAlaTyrAlaLysGlu	566	
Db	486105	GATGATGTTGAAAATCAACAAAAAGTT	ATTATTCTGAACCTCAAGTAGAAC	CAACAACTCGGA	486046
Qy	567	LysGlyLeuThrProProSerThrAspHis	GlnAspSerGlyAsnThrGluAla	584	
Db	486045	GAAACAGTTTTTGAACCATCTGCAAGAG	CTAAATTTGATTCACAGTTGAATCG	GTTCAA	485986
Qy	585	---LysGlyAlaGluAlaIleTyrAsnArg	ValLysAlaAlaLysLysVal---ProLeu	602	
Db	485985	GATTCACAACTGAACTGTCTCTTGAAGA	AGTTCAAACCTCAACAGAAATTC	AACTGTA	485926
Qy	603	AspArgMetProTyrAsnLeuGlnTyrThr	ValGluValLys-----	616	
Db	485925	GAACTCTCAACAGAAAGCTACTCTTTGA	TACTGTTTCAACTGACAAAATCTCT	CAAGAAAGCT	485866

APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 145:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10711 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-145

Alignment Scores:
 Pred. No.: 0.00233 Length: 10711
 Score: 151.00 Matches: 170
 Percent Similarity: 32.92% Conservative: 119
 Best Local Similarity: 19.36% Mismatches: 340
 Query Match: 3.76% Indels: 249
 DB: 4 Gaps: 42

JS-09-765-272A-66 (1-763) x US-08-961-527-145 (1-10711)

QY	5	LeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSerTyrIleAsp	24
DB	444	TTGGATGCTATGTAAGGGGGAATCAAA-----GAT	476
QY	25	GlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGluValSerLys	42
DB	477	TTCTCACAGTAGGAATTAAAGGCTCTGGATGAACACACAGAGTTCAGTACACTTTGAACAAA	536
QY	43	ArgGluGlyIle-----AsnAlaGluGln***ValIleLysIleThrAspGlnGly	59
DB	537	CCAGAAAGCTTCGGAAATTCCTAAGACAACATCGGGTGTCTTCGCGCCAGCTTAATGAAGAG	596
QY	60	TyrValThrSerHisGlyAspHisTyrHis-----TyrTyr	71
DB	597	TTTTTGAATTCAAAGGAGATGATTTTGGCAAGCTACGGATCCAAAGTAGTAGTCTCTGTAT	656
QY	72	AsnGlyLysValProTyrAspAlaIleIleSerGluGluLeu-----MetLysAsp	89
DB	657	AACGGTCTCTATTGTTGTAATCCATTGTGACCAAAATCCTCTGTGAAATTCGGAATAAT	716
QY	90	ProAsnTyrGlnLeuLysAspSerAspIleValAsnGluIleLysGlyTyrValIle	109
DB	717	CCGAACACTCTGGGTAAGACAAAT-----GTGCATGTTGAC	752
QY	110	LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg	129
DB	753	AAAGTTAAATGTCATCTCTGGATGCTCAAGATACAGCAACCTGCAGAAAACCTTTAAA	812
QY	130	-----ThrLysGluGluIleLysArgGlnLysGlnGluArgSerHisAsnHis	145
DB	813	GATGGTAGCCTTACAGCAGCTGCTCTCTATCCAAACAAAGTGCAGTTTTCGCAGAACTTGAG	872
QY	146	AsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAsp	165
DB	873	AAAGAGTATGAAGCACAAATATGTC-----TATACTCAACAA	908
QY	166	AspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleVal	185
DB	909	GACTCTATTACGTATTAGTTGGTACAAATATTGACCGCTCAGTCCTTATAAATACATCT	968
QY	186	ProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAla	205
DB	969	AAGACCAGGACGACAAAGGATCGACTAAAGAGCTCTCTTAAACAAAGGATTTCCGT	1028
QY	206	AlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyr	225

```
QY 539 ProHisMetThrHisSerHisTriPileLysLysAspSerLeuSerGluAlaGluArgAla 558
Db 1908 TACCAAAAAGCTCAGGAAAAATGGATGAAGAAAA-----GAAGAGTCTTAATAAAAAG 1961
QY 559 AlaAlaGlnAlaTyraLys-----GluLysGlyLeu 569
Db 1962 GCTCAAGAAGATCTCGCAAAACATGTGAAATAACTGTGCAAAATATAAGAAAGATTTA 2021
QY 570 ThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAla 589
Db 2022 GTATTTCCCTTG-----AATGCTGAATCCCTTTTATCAATTTGTAAGAAAGATCTTAAA 2075
QY 590 IleTyraAsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyraAsnLeu 609
Db 2076 ATGTACGACGCCCAAGAGTTGGAGCCTCTTTTGTGCAGATAGAGAAATTTTGTGTTAA 2135
QY 610 GlnTyraThrValGluValLysAsnGly-----SerLeuIleLe 622
Db 2136 TTTTACTTGTTCCTATTTGCTTTCTCAGCTATTATTGTTATATTATAAAGTATAATTATT 2195
QY 623 ProHis-TyraSp-----HisTyraHisAsnIleLysPhe----- 633
Db 2196 TTTTATTATCAGAGTTAAGCATTCGACTTTTCAGAGGAGGAGTATTTTAAAAAGAAA 2255
QY 634 -----GluTyrPheAspGluGlyLeuTyraG 642
Db 2256 ATGTAAACGTTTGTCTCAAAATGAAGGATTAGAGTTTATGAATAAAGGATTATTGTA 2315
QY 642 uAlaProLysGlyTyraThrLeuGlu-----AspLeuLeuAlaTh 655
Db 2316 AAAACGTTGTAAATATAGTATTTCGAAATTTTCATTAGTGTGTGCTTCTGTTATGATTGG 2375
QY 655 rValLysTyraTyraValGluHisProAsnGluArgProHisSerAspAsnGly----- 672
Db 2376 AGTGCATTCTTTGGGCAAGTCCGGTCTTTCAGATAGCAGTGCAGTCTGGTTCACCGC 2435
QY 673 -----PheGlyAsnAlaSerAspHisValGlnArgAs 683
Db 2436 GAACTTACCAGCTAGTTAGCTACTGCTCTTGCACACGAAAGAGATGATGGCGGTGA 2495
QY 683 n-----LysAsnGly-----GlnAlaAspThrAsnGlnThrGluLysProSe 697
Db 2496 TTTTGAAGCGCTAAGTGGGAGAGACCAAGTTCTCCAGAAAGTTACAGATGGACCTAA 2555
QY 697 rGluLysPro-----GlnThrGluLysProGluGluGluThrProArgG 713
Db 2556 GACAGAAGAAGAACTATTAGCACTTGAAAGAAAGAAACCGGCTGAAGAAAGAAACCAAGA 2615
QY 713 uGluLysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSe 733
Db 2616 GGATTAACCTCGACGTCTAAACCTGAACCTGAACACCTTAAG---ACGGTAACCCCTGAATGGA 2672
QY 733 rProGluGluSerGluGluProGlnValGluThrGluLysValGluGluLys 750
Db 2673 AACGGTAGCGAATAAGAGCAACAGGGAACAGTCACTATCCGAGAGAGAAAA 2724
```

RESULT 38

```
US-09-286-981B-26
; Sequence 26, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenmann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 26
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-26
```

```
Alignment Scores:
Pred. No.: 9,35e-05 Length: 1338
Score: 150.50 Matches: 126
Percent Similarity: 34.9% Conservative: 72
Best Local Similarity: 22.2% Mismatches: 213
Query Match: 3.74% Indels: 156
DB: 4 Gaps: 26
```

US-09-765-272A-66 (1-763) x US-09-286-981B-26 (1-1338)

```
QY 214 LysGlnGlySerArgProSerSerSerSerSerTyraAsnAlaAsnProAlaGlnProAr 233
Db 6 GAACGAGGAGGAGTACCAGAGTACCACCTCTTCTAATAGGCAATGAAAGTCAGGCAGA 65
QY 233 gLeuSerGluAsnHisAsnLeu-----ThrValThrProThrTyraHisGlnAsnGln 251
Db 66 ACAAGGAGAACCAACCTAAAAAACTCGAATTCAGAACGAGATAAGCGCAAGGAAGAGGTGCA 125
QY 251 yGluAsnIleSerSerLeuLeuArgGluLeuTyraLysProLeuSerGluArgHisVa 271
Db 126 GGAATATGTAATAAATAATAGTGGTGAGAGCTATGCAAAATCAATCAATAAAGGACATAC 185
QY 271 lGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyVa 291
Db 186 AATTACTGTAGCTCTAGTTAACGAGTTGACACATTAG----- 225
QY 291 lAlaValProHisGlyAsnHisTyraHisPheLeuProTyraGluGlnMetSerGluLeu 311
Db 226 -----AACGAGTATTGTAATAAATAGTTGAATCAACCTCAGAAAGCCA 269
QY 311 uLysArgIleAlaArgIleLeuProLeuArgTyraArgSer-----As 325
Db 270 A-----CTACAGATATCTGATGATGGAGAGTCCGATCAAAAGTAGATGAGTGTGTC 320
QY 325 nHisTyrValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSe 345
Db 321 TAAATTGGAAGAGGACTCATCTCTTCGTCAGTCTCCACT---AAACCGGA 377
QY 345 rProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProLysAspGlu-- 364
Db 378 AGCTTCAGATACAGCGAAGCCAAACAAAGCCGACAGAACCAAGGAGAGGAGGTAGCAGAAC 437
QY 365 ----LysLeuValLysGluAlaValArgLysValGlyAspGlyTyraValPheGluGluAs 383
Db 438 TAAAGAAGAGGTTGAAGAAGCTGAGAAAGAAAGCCAAAGGATCAA-----AAAGAGA 488
QY 383 nGlyValSerArgTyra-----IleProAlaLysAspLeuSerAlaGluThrAla---Al 400
Db 489 AGATCGTCTTAATCTACCACCTTACTTACAAACGCTTGAACTTGAAATTCGTGAGTC 548
QY 400 aGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLy 420
Db 549 CGATGTGGAAGTTAAAGAAAGCGAGCTGAA-----CTAGTAAAGGTGAA 593
QY 420 sThrAspLeuProSerSerAspArgGluPheTyraAsnLysAlaTyraAspLeuLeuAlaAr 440
Db 594 AGCTAACGAACTCGAGACGACGACCAAAATTT----- 624
QY 440 gIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAs 460
Db 625 -----AAGCAAGCAGAGAGCGAGAGTTCAGAGTAAACAGC 659
QY 460 nLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLe 480
```



```
660 TGAGCTACAGGTTTAAAAAATCAAGACACATCTCTGAAGAGCAGACAGAA----- 714
480 uAlaPheLeuAlaProIleArg---HisProGluArgLeuGlyLeuGlyProAsnAlaGlnI 499
715 -----GCTAAACGAGCAGAGTGTCTTAAGAGCAGAGTAAACCAAGGGGGG--- 762
499 eThrTyrThrAspAspGluIleGlnValAlaLys-----LeuAlaGlyLysTyrThr 517
763 -----GCAAAACGAGGAGTCTCTGGAGAGCTAGCAAC 794
517 rGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVa 537
795 ACCTGAT----- 801
537 lThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluAr 557
802 -----AAAAA-----GAAATGATGC 818
557 gAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisG 577
819 GAAGCTCTCAGATTCTAGGTAGGTGAAGAACTCTTCCAGCCCATCCCTG----- 870
577 nAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAl 597
871 -----AAACCCAGAAAAAAGGTAGCAGAGCT---GAGAAGAAGGTTGAGAAGC 917
597 aLysLysValProLeuAspArg-----MetProTyrAsnLeuG 610
918 TAAGAAAAAAGCCGAGGATCAAAAAAGAGAGATCCCGTAACCTACCAACCAACTT 977
610 nTyrThr-----ValGluValLysAsnGlySerLeuI 621
978 CAAACGCTTGAACTTGAATTCCTGAGTCCGATGTTGAAGTTAAATAAAGCGAGCTT-- 1035
621 eIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTy 641
1036 -----GAACTAGTATAA 1046
641 rGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValG 661
1047 AGAGGAAGCTAAGGAACCTCGAAACCGAGGAAAGTTAAGCAAGCAAAACCGGAAGTTGA 1106
661 uHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValG 681
1107 G-----AGTAAAAAGCTGAGGCTACAGGTAGTGAAGTAAATAAATCAA 1145
681 nArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysPr 701
1146 GACAGATCTATAA---AAAGCAGAGAAGAGCTAAACGAAAGACGAGCAGAGAAGATAA 1202
701 oGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLysPr 721
1203 AGTTAAAGAAAAAACCAAGCTGAACAACACCAACACGAGCGCGCTCCAAAAGCAGAAAAAC 1262
721 o---GluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPr 740
1263 AGCTCAGCTCCAAAACCA---GAGAATCCAGCTGAACAACCAACCAAGCAGAAAAACCA 1319
740 oGlnValGluThrGluLys 746
1320 TGATCAACAGCTGAAGAA 1338

RESULT 39
US-09-535-008-66
; Sequence 66, Application US/09535008
; Patent No. 645629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 2318-259
```

```
; CURRENT APPLICATION NUMBER: US/09/535,008
; EARLIER FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 66
; LENGTH: 5468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (75)..(5012)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (1780)
; OTHER INFORMATION: GenBank Accession No. 645629 U29175 shows a C at this
; OTHER INFORMATION: position (position 1784 in GenBank) rather than
; OTHER INFORMATION: the G shown here.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (26)
; OTHER INFORMATION: Polymorphism of either T or C in this noncoding
; OTHER INFORMATION: region.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1583)
; OTHER INFORMATION: Polymorphism of A or G resulting in a silent
; OTHER INFORMATION: mutation.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1598)
; OTHER INFORMATION: Polymorphism of T or C resulting in a silent
; OTHER INFORMATION: mutation.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1892)
; OTHER INFORMATION: Polymorphism of A or G resulting in a silent
; OTHER INFORMATION: mutation.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4498)..(4499)
; OTHER INFORMATION: A CAG is missing between these bases as compared
; OTHER INFORMATION: to SEQ ID NO:1.
; US-09-535-008-66
```

```
Alignment Scores:
Pred. No.: 0.00887 Length: 5468
Score: 150.50 Matches: 146
Percent Similarity: 30.35% Conservative: 81
Best Local Similarity: 19.52% Mismatches: 241
Query Match: 3.74% Indels: 280
DB: 4 Gaps: 31
```

US-09-765-272A-66 (1-763) x US-09-535-008-66 (1-5468)

```
Qy 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
Db 353 GAAAGGAATGGGATGCGGTCTCAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCCCAT 412
Qy 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAl 207
Db 413 GGACCAGCACTCCCAAGGTTTACCCC---TCGCCCCCTGGGTGGCTCTGACATGCCCTCTAG 469
Qy 207 aGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsnAl 227
Db 470 TCCAGTT-----CCAGCAGTGGCCCGTCTCGGGGCCCCCAGATGTCCTTC 514
Qy 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240
Db 515 CGGGCCAGGAGTGCCCGCTGGATGTGCTGACCCCGAGGCTTGGGGCAGCAGAACCG 574
Qy 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 620
```

Db 575 GGGCCCAACCCCA---TTTAAACGAGAACCCAG-----CTGCACCA 610
QY 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280
Db 611 GCTCAGAGCTCAGATCATGCTCAAGATGCTG----- 644
QY 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
Db 645 -----GCCAGGGGCGAGCCCTCCCGACACCTGCAGATGGC 682
QY 300 sPheIleProTyrGlnGlnMetSerGluLeuGluLysArgIleAlaArgIle----- 317
Db 683 GGTGCAGGGCAAGCGCGCATGCTCCGGATGCAGCAGCAGATGCCAAGCTTACCTCCACC 742
QY 317 ----- 317
Db 743 CTCGGTGTCCGCAACAGACCGCGCCCTGGCCCTGGCCCTCCCGCCCGCGGTCGCCG 802
QY 318 -----IleProLeuArgTyrArgSerAsnHis----- 326
Db 803 CCGCGCACCTCCAAATTCAGCAGGCTCATGTTATGGAGGGCCCAACATGCCTTCCCC 862
QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336
Db 863 AGGACCTCGGGCGTGCCCCCGGGATGCCAGGCCAGCTCTCGAGGGCCCTCCCAAGCC 922
QY 337 -----SerProGlnSerThrProGlu----- 343
Db 923 CTGGCTTGAAGGACCCATGTGCGAATGTGTCCTCCCGCCACGACGCCCTCAGAAGCTGAT 982
QY 344 -----ProSerProSerProGlnProAlaProAsnPro----- 354
Db 983 TCCCCCGACGACCAAGCGCGCCCTTCCCCCGCGCCCTCCGCTCCACCGCGGCGCTC 1042
QY 355 -----GlnProAlaProSerAs 360
Db 1043 GCCCGTGATCCACCGCAGACCCAGTCCCGCGGCGAGCGCGCCCGCCCGCCATGTT 1102
QY 360 nProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
Db 1103 GCCACTGCACCAAGCAGAGCGCGCATCACCCCATCCAGAAAGCGCGGGCGCTCGACCC 1162
QY 378 -----TyrValPheGluGlnAsnGlyValSerArgTyrIl 389
Db 1163 TGTGGAGATCCTGCAGAGCGAGTACAGGCTCGAGCTCGCATCGCACCCGAATTC 1222
QY 389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405
Db 1223 GGAACCTGAAACCTTCCCGGTCCCTGCGCGGGATTTGCGAACCAAGCGACCATGA 1282
QY 405 ----- 405
Db 1283 GCTCAAGGCCCTCAGGCTGTGAACCTCCAGAGCGAGCTGGCGCAGGAGTGTGGTGTG 1342
QY 406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuLysLeuLysLeuLysLeuLysLeuLys 423
Db 1343 CATCGGAGGGACACAGCGCTGGAGACAGCCCTCAATGCTAAGCGCTTACAGCGCGAGCAA 1402
QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGl 443
Db 1403 CGCGCATGCTCCGCGAG-----GCCGATCACTGA 1435
QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheGl 456
Db 1436 GAACCTGGAGAGCAGCAGAGATCGACAGAGCGCGCAAGCGCGCAGACCAAGCA 1495
QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
Db 1496 ATACCTCAATGACATTCACAGCATGCCAAGGATTCACAGGAATATCACAGATCCGTCAC 1555
QY 469 *SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle----- 486

Db 1556 AGGCAAAATCCAGAGCTGACCAAGAGGAGTGGCCACGTACCATGCCAACACGAGGCGGA 1615
QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
Db 1616 GCAGAGAAGAGAGACGAGCGATCGAGAGAGCGCATGCGGAGCTCATGCTGAAGA 1675
QY 504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAs 524
Db 1676 TGAGGAGGGGTACCCCAAGCTCATC----- 1700
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHis 544
Db 1701 -----GACCAGAGAAGAGCAAGCGCTGCCCTACCTCTTG-----CAGCAGACAGCA 1750
QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAl 564
Db 1751 GTACGTGCTAACCTCAGGAGCTGGTGGGCGAGCAGACAGGCTGCCAGGTCCCAAGGA 1810
QY 564 aLysGluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
Db 1811 GAAAAG 1867
QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604
Db 1868 CATTCGGCGGATGGC-----GAGCCTCTAGACGA 1897
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGl 618
Db 1898 GACCAACGAGATGAGCGACCTCCCGTGAAGTG-----ATCCAGTGGAGAGTGG 1948
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGl 638
Db 1949 GAAGATCTCCACAGC----- 1964
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658
Db 1965 -----ACAGATGCCCCCAAGCGCGGCGAGCTGGAGGCTGGCTCGAGATGAACCGGG 2017
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
Db 2018 GTAT-----GAAGTAGCTCCGAGGTCTGAT----- 2042
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGl 698
Db 2043 -----AGTGA 2047
QY 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
Db 2048 AGAAAGTGGCTCAGAGAAG 2101
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGl 738
Db 2102 AGCAGAGCTCCACCC-----CTGCCGTGGAGAGAGAGAGAGATTCAGATCCAGACAG 2158
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaGl 755
Db 2159 CGATCAGCTCTCTGAGTGGAGCGCGCGGCACATCATTTGAGATGCCCAAGCAAGATGCGA 2218
QY 755 uAspLeuLeuGlyLysIleGln 762
Db 2219 TGATGAATATGCGTGTCCAG 2240
RESULT 40
US-09-535-008-1
; Sequence 1, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 2318-259

/ CURRENT APPLICATION NUMBER: US/09/535,008
 / CURRENT FILING DATE: 2000-03-23
 / EARLIER APPLICATION NUMBER: U.S. 60/125,806
 / EARLIER FILING DATE: 1999-03-23
 / NUMBER OF SEQ ID NOS: 77
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 1
 / LENGTH: 5471
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (75)..(5015)
 / FEATURE:
 / NAME/KEY: conflict
 / LOCATION: (1780)
 / OTHER INFORMATION: GenBank Accession No. 6465629 U29175 shows a C at this
 / OTHER INFORMATION: position (position 1784 in GenBank) rather than
 / OTHER INFORMATION: the G shown here.

/ FEATURE:
 / NAME/KEY: allele
 / LOCATION: (26)
 / OTHER INFORMATION: Polymorphism of either T or C in this noncoding
 / OTHER INFORMATION: region.

/ FEATURE:
 / NAME/KEY: allele
 / LOCATION: (1583)
 / OTHER INFORMATION: Polymorphism of A or G resulting in a silent
 / OTHER INFORMATION: mutation.

/ FEATURE:
 / NAME/KEY: allele
 / LOCATION: (1598)
 / OTHER INFORMATION: Polymorphism of T or C resulting in a silent
 / OTHER INFORMATION: mutation.

/ FEATURE:
 / NAME/KEY: allele
 / LOCATION: (1892)
 / OTHER INFORMATION: Polymorphism of A or G resulting in a silent
 / OTHER INFORMATION: mutation.

JS-09-535-008-1

Alignment Scores:
 Pred. No.: 0.000888 Length: 5471
 Score: 150.50 Matches: 146
 Percent Similarity: 30.35% Conservative: 81
 Best Local Similarity: 19.52% Mismatches: 241
 Query Match: 3.74% Indels: 280
 Gaps: 31

JS-09-765-272A-66 (1-763) x US-09-535-008-1 (1-5471)
 2Y 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
 353 GAAAGGAATGGGATGGCGTGCAGGGGCCATGCTGGGATGGGGCCCGCCAGCCCAT 412
 2Y 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAl 207
 413 GGACCAAGCATCCCAAGGTATACCC---TCGCCCTGGGTGGCTCTGAGCATGCTCTAG 469
 2Y 207 aGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl 227
 470 TCCAGTT-----CCAGCATGGCCGCTCTTCGGGGCCCGAGATGCTTC 514
 2Y 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240
 515 CGGGCCAGGAGGTGCCCGCTGGATGGTGTGACCCCGAGCCTTGGGGCCCGAGAACCG 574
 2Y 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 260
 575 GGCCCCAACCCCA---TTTAACCAAGAACCCAG-----CTGCACCA 610
 2Y 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280

Db 611 GCTCAGAGCTCAGATCATGGCCCTACAAGATGCTG----- 644
 QY 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
 Db 645 -----GCCAGGGGCGAGCCCTCCCGCCACCTGCAGATGGC 682
 QY 300 sPheIleProTyrGluGlnMetSerGluLeuGlnLysArgIleAlaArgIle----- 317
 Db 683 GGTGCAGGGCAAGCGCGGATGCCCGGGATGCAGCAGCAGATGCCAACGCTACCTCCACC 742
 QY 317 ----- 317
 Db 743 CTCGGTGTCCGCAACAGGACCCGCGCTGGCCCTGGCCCTGGCCCGCGCGGTCCCGG 802
 QY 318 ----IleProLeuArgTyrArgSerAsnHis----- 326
 Db 803 CCCGGCAGCTCCAATTACAGCAGGCGCTCATGGTATGGAGGGCCCAACATGCTCCCC 862
 QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336
 Db 863 AGGACCTCGGGCGTGCCTCCCGGATGCCAGGCCAGCCTCTCTGGAGGGCTCCCAAGCC 922
 QY 337 -----SerProGlnSerThrProGlu----- 343
 Db 923 CTGGCCTGAAGACCCATGGCGAATGTGCTGCCCCACGAGCAGCCCTCAGAAGCTGAT 982
 QY 344 ---ProSerProSerProGlnProAlaProAsnPro----- 354
 Db 983 TCCCCCGCAGCCAGCGGCGCCCTTCCCGCGGCCCTGCTGCCCTCCACCCGCGCCTC 1042
 QY 355 -----GlnProAlaProSerAs 360
 Db 1043 GCCCGTGTATGCCACCGCAGACCCAGTCCCGCGGCGAGCCGCCCGCCCGCCATGGT 1102
 QY 360 nProIleAspGlnLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
 Db 1103 GCCACTGCACCAAGACGAGCGCGCATCCCCCATCCAGAGCCCGCGGGGCTCGACCC 1162
 QY 378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389
 Db 1163 TGTGGAGATCTCGCAGGAGCGCGAGTACAGGTGTCAGGCTCGCATCGCACCGCAATTCA 1222
 QY 389 eProAlaLysGlnGluSerLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405
 Db 1223 GGAACCTTGAACACCTTCCGGGTCCCTGCGCGGGATTTGCGNACCAAGACGACCATGA 1282
 QY 405 ----- 405
 Db 1283 GCTCAAGGCCCTCAGGCTGTGAACCTTCAGAGGCGAGCTGCGCCAGGAGGTGGTGTG 1342
 QY 406 -LeuAlaLysGlnGluSerLeuSerHisLysIleGlyAlaLys-----LysThrAspLe 423
 Db 1343 CATCGGAGGACACAGCGCTGAGACAGCCCTCATGCTAAGGCTTACAAGCGCAGCAA 1402
 QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisG 443
 Db 1403 CGCCAGTCCCTGCCGAG-----GCCCGCATCACTGA 1435
 QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456
 Db 1436 GAAGCTGGAGAAGCAGCAGAGATCGAGCAGCGCAGCGCCGCGCAGAGCAGCAGGA 1495
 QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
 Db 1496 ATACCTCAATGACATCTCCAGCATGCCAAGGATTTCAAGGAATATCAGATCCGTCAC 1555
 QY 469 *SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle----- 486
 Db 1556 AGGCMAAATCCAGAGCTGACCAAGCGAGTGCAGGCTACCATGTCACCAACGAGCGGGA 1615
 QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
 Db 1616 GCAGAGAAAGNAGACGAGCGGATCGAGNAGGCGCATCGCGAGGCTCATGCTGAAGA 1675

```
QY 504 pGluileGlnValAlaLysLeuAlaGlyLysTyrThrThruAspGlyTyrIlePheAs 524
Db 1676 TGAGGAGGGTACCGCAAGCTCATC-----1700
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544
Db 1701 -----GACCAGAAGAAGCAGCGCTGGCCCTCTCTTG---CAGCAGACAGCAGA 1750
QY 544 rHisTIPleIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564
Db 1751 GTACGTGGTAACTCAGGAGCTGGTGGCGCACACAAAGCTGCCAGTGGCCCAAGGA 1810
QY 564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
Db 1811 GAAAAAGAGAAAAAGAAAAAGAGAGGAGCAAAATGCAGAA---GGACAGACGCTGC 1867
QY 584 aLysGlyAlaGluAlaIleTyrIleArgValLysAlaLysLysValProLeuAspAr 604
Db 1868 CATTGGGCGGATGCG-----GAGCCTCTAGACGA 1897
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnG1 618
Db 1898 GACCAGCCAGATGAGCGACCTCCCGGTGAAGTG-----ATCCACGTGGAGAGTGG 1948
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspG1 638
Db 1949 GAAGATCTCACAGC-----1964
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658
Db 1965 -----ACAGATCCCCAAAGACCGGCGCAGCTGGAGCCTGGCTCGAGATGAACCCGG 2017
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
Db 2018 GTAT-----GAAGTAGCTCCGAGGTCTGAT-----2042
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG1 698
Db 2043 -----AGTGA 2047
QY 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
Db 2048 AGAAGTGGCTCAGACAGAGAGAGAGAGAGAG-----GAGGAAGAGCAGCGCCAGGC 2101
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG1 738
Db 2102 AGCACAGCTCCACC---CTGCCCGTGGAGGAGAGAGAAAGATTCCAGATCCAGACAG 2158
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaG1 755
Db 2159 CGATGAGCTCTGAGGTGGACCGCGCGGCACATCATTTAGAAATGCCAAGAAATGTGCA 2218
QY 755 uAspLeuLeuGlyLysIleGln 762
Db 2219 TGAATGATATGGCTGTCTCCAG 2240
```

RESULT 41

```
US-09-535-008-74
; Sequence 74, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 2318-259
; CURRENT APPLICATION NUMBER: US/09/535,008
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 77
```

Alignment Scores:

Pred. No.:	0.000889	Length:	5477
Score:	150.50	Matches:	146
Percent Similarity:	30.35%	Conservative:	81
Best Local Similarity:	19.52%	Mismatches:	241
Query Match:	3.74%	Indels:	280
DB:	4	Gaps:	31

US-09-765-272A-66 (1-763) x US-09-535-008-74 (1-5477)

```
QY 177 GluAspThrGlyAspAlaTyr-----lIleValProHis 187
Db 353 GAAAGGAATGGGGATGCGGTACAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCCAT 412
QY 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207
Db 413 GGACCAGCAGCTCCCAAGGTTACCCC---TCGCCCTTGGGTGGCTCTGAGCATGCTCTAG 469
QY 207 aGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl 227
Db 470 TCCAGTT-----CCAGCCAGTGGCCGCTCTCGGGGCCCCAGATGCTTC 514
QY 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisLe 240
Db 515 CGGGCCAGGAGGTGCCCGCTGGATGTTGTCACCCCGGCTTGGGGCAGCAGAACCG 574
```


FILE REFERENCE: 2318-259
CURRENT APPLICATION NUMBER: US/09/535,008
CURRENT FILING DATE: 2000-03-23
EARLIER APPLICATION NUMBER: U.S. 60/125,806
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 70
LENGTH: 5480
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (75)..(5024)
FEATURE:
NAME/KEY: conflict
LOCATION: (1780)
OTHER INFORMATION: GenBank Accession No. 6465629 U29175 shows a C at this position (position 1784 in GenBank) rather than the G shown here.
OTHER INFORMATION: Polymorphism of either T or C in this noncoding region.
FEATURE:
NAME/KEY: allele
LOCATION: (26)
OTHER INFORMATION: Polymorphism of A or G resulting in a silent mutation.
FEATURE:
NAME/KEY: allele
LOCATION: (1583)
OTHER INFORMATION: Polymorphism of T or C resulting in a silent mutation.
FEATURE:
NAME/KEY: allele
LOCATION: (1598)
OTHER INFORMATION: Polymorphism of T or C resulting in a silent mutation.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4245)..(4253)
OTHER INFORMATION: Insertion of 9 basepairs as compared to SEQ ID NO:1.
OTHER INFORMATION: No:1.

US-09-535-008-70

Alignment Scores:
Pred. No.: 0.00089 Length: 5480
Score: 150.50 Matches: 146
Percent Similarity: 30.35% Conservative: 81
Best Local Similarity: 19.52% Mismatches: 241
Query Match: 3.74% Indels: 280
DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-70 (1-5480)

177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
353 GAAAGGAATGGGGATGCGGTGAGGGGGCCATGCTGGGATGGGGCCCGCCCGCCCGCCCAT 412
188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207
413 GGACGAGCACCACAGGTTACCC---TCGCCCCCTGGGTGCTCTGAGCATGCTCTAG 469
207 aGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl 227
470 TCCAGTT-----CCAGCCAGTGGCCGGTCTTCGGGGCCCGCCAGATGCTTC 514
227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240
515 CGGGCCAGAGGTGCCCGCTGGATGGTGTGCTGACCCCGCCAGGCTTGGGGCGCAGAACCG 574

240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 260
575 GGGCCCAACCCCA---TTTAAACAGAACCA---CTGCACCA 610
260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280
611 GCTCAGAGCTCAGATCATGGCTACAGATGCTG--- 644
280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
645 -----GCCAGGGGGAGCCCTCCCGCAGCCACCATGTCAGATGGC 682
300 sPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIle----- 317
683 GGTGAGGGCAAGCGCGCATGCGCGGATGCGAGCAGCAGATGCCAACGCTACCTCCACC 742
317 ----- 317
743 CTCGGTGTCCGCAACAGGACCCCGCCTGGCCCTGGCCCTGGCCCGCCCGGGTCCGG 802
318 ---IleProLeuArgTyrArgSerAsnHis----- 326
803 CCGGCACCTCCAAATTACAGCAGGCTCATGTATGGAGGGCCCAACATGCTCCCCC 862
327 -----TrpValProAspSerArgProGluGlnPro----- 336
863 AGGACCTCGGGCGTGCCTCCCGGATGCCAGGCAGCTCTCTGGAGGGCTCCCAAGCC 922
337 -----SerProGlnSerThrProGlu----- 343
923 CTGGCTGAAGGACCCATGGCGAATGCTGTCGCCCGCCAGCAGCACCCTCAGAAGTGTAT 982
344 ---ProSerProSerProGlnProAlaProAsnPro----- 354
983 TCCCCCGCAGCAACGGCGCGCTTCCCCCGGCCCTCCCGCGGCCCTGCCCTCCACC 1042
355 -----GlnProAlaProSerAs 360
1043 GCCGTGTATGCCACCGCAGACCCAGTCCCGCGGCAGCCGCGCCGCGCCCATGGT 1102
360 nProIleAspGlyLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
1103 GCCACTGCACCAAGCAGCAGCGCGCATACCCCATCCAGAACGCCGGGGCTCGACCC 1162
378 -----TyrValPheGluAsnGlyValSerArgTyrI 389
1163 TGTGGAGATCTGCAGAGCGCGAGTACAGGCTGCAGGCTGCATCGCACACCAATCA 1222
389 eProLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405
1223 GGAACCTTGAAACCTTCCCGGGTCCCTGGCGGGGATTTCCGAAACCAAGCGACCATGA 1282
405 ----- 405
1283 GCTCAAGGCCCTCAGGCTGCTGAACCTTCCAGAGCAGCTCGCCGAGGAGTGTGGTGTG 1342
406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423
1343 CATGGAGGAGGACAGACCGCTGGAGCAGCCCTCAATGCTAAGCCTTACAGCGCAGCAA 1402
423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisG 443
1403 GCGCCAGTCCCTGGCGAG-----GCCCGCATCACTGA 1435
443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456
1436 GAAGCTGGAGAGCAGCAGAGATCGAGCAGGAGCGCAGCGCGCAGAGCAGCAGGA 1495
456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
1496 ATACCTCAATGATCTCTCCAGCATGCCAGGATTTCAAGAAATATCAGAGATCCGTCAC 1555
469 *SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle----- 486

470 TCCAGTT-----CCAGCCAGTGGCCGCTCTCGGGGCCCCAGATGCTCTTC 514
227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisLeu 240
515 CGGCGCAGGAGTCCCGCTGGATGGTGTGACCCCGAGGCTTGGGGCAGCAACCG 574
240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnLeuSerSerLeuLeuArgGln 260
575 GGGCCCAACCCCA---TTTAACCAAGAACAG-----CTCCACCA 610
260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuLeuPheAspPr 280
611 GCTCAGAGCTCAGATCAGGCTACAGATGCTG----- 644
280 oAlaGlnInleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
645 -----GCCAGGGGCGAGCCCTCCCGACCACTCCAGATGGC 682
300 sPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle----- 317
683 GGTGAGGCAAGCGCGATGCGCGGATGCGAGCAGATGCCAACGCTACCTCCACC 742
317 ----- 317
743 CTGGTGTCCCAACAGACCCGGCCCTGGCCCTGGCCCTGGCCCGCGGGTCCCGG 802
318 -----IleProLeuArgTyrArgSerAsnHis----- 326
803 CCGGGCACCTTCAAAATTACAGCAGCCTCATGGTATGGAGGGCCCAACATGCTCCGCC 862
327 -----TrpValProAspSerArgProGluGlnPro----- 336
863 AGGACCCTCGGGCGTGGCCCGCGGATGCCAGCGCAGCTCTCTGAGGGCCCTCCCAAGCC 922
337 -----SerProGlnSerThrProGlu----- 343
923 CTGGCTTGAGGACCCATGGCGAATGTGCTGCTCCCGCCACAGCACCCCTTCAGAACTGAT 982
344 -----ProSerProSerProGlnProAlaProAsnPro----- 354
983 TCCCGCGCAGCAAGCGGCGCCCTTCCCGCGCGCCCTGCGTCCCGCCCGCGCCTC 1042
355 -----GlnProAlaProSerAs----- 360
1043 GCCCGTGATGCCACCGCAGACCCAGTCCCGCGGCGAGCGCGCCAGCGCGCGCCATGT 1102
360 nProIleAspGlyLeuValLysGluAlaValArgLysValGlyAspGly----- 377
1103 GCCACTGACCAAGAGCAGCGCGCATCACCCCGCATCCAGAGCGCGGGGCTCGACCC 1162
378 -----TyrValPheGluGluAsnGlyValSerArgTyrIle 389
1163 TGTGGAGATCTCGCAGGAGCGGAGTACAGGCTGCGAGCTCGCATCGCACCCGAATCA 1222
389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405
1223 GGAACCTTGAAACCTTCCCGGGTCCCTGCGCGGATTTGCAACCAAGCGACCATCA 1282
405 ----- 405
1283 GCTCAAGGCCCTCAGGTGTGAATCTCCAGAGGCGAGTGGCCCGAGGAGGTGGTGGTGG 1342
406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423
1343 CATCGGGAGGACACAGCGCTGGAGACAGCCCTCAATGTAGGCTACACAGCGCAGCA 1402
423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGln 443
1403 GCGCCAGTCCCTGCGCGAG-----GCCCGCATCACTGA 1435
443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheGln 456
1436 GAAGCTGGAGAGCAGCAGAGATCGAGCAGGCGGCAAGCGCCGCGCAGACCAAGCA 1495

RESULT 44

US-09-535-008-64

; Sequence 64, Application US/09535008

; Patent No. 6465629

456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
1496 ATACCTCAATAGCATCTCCAGCATCCCAAGGATTTCAAGGAATATACAGATCCGTCAC 1555
469 *SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle----- 486
1556 AGCAAAATCCAGAACTGACCAAGCAGTGGCGCAGTACCATCGCAACCGGAGCGGA 1615
487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
1616 GCAGAGAAAGAGAACGAGCGATCGAAGAGGAGCGCATCGGAGGCTCATGGCTGAAGA 1675
504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524
1676 TGAGGAGGGGTACCGCAAGCTATC----- 1700
524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544
1701 -----GACCAAGAAAGACCAAGCGCTCGGCTACCTCTTG---CAGCAGCAGAGGA 1750
544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564
1751 GTACGTGGCTAACCTCAGGAGCTGGTGGCGCAGCACAGGCTGCCAGGTCCCAAGGA 1810
564 aLysGluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
1811 GAAAAGAGAAAGAAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1867
584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604
1868 CATTGGCGCGGATGGC-----GAGCCTCTAGACGA 1897
604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGln 618
1898 GACCACCGCAGATGAGCGACTCCCGGTGAGGTG-----ATCCAGCTGGAGAGTGG 1948
618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGln 638
1949 GAAGATCTCTCACAGG----- 1964
638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTy 658
1965 -----ACAGATGCCCCCAAGCCCGGCGAGGCTGGCTCGAGATGAACCCGGG 2017
658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
2018 GTAT-----GAAGTAGCTCCGAGTCTGAT----- 2042
678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGln 698
2043 -----AGTGA 2047
698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
2048 AGAAAGTGGCTCAGAAAG 2101
718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGln 738
2102 AGCACAGCTTCCACC-----CTGCCCTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2158
738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaGln 755
2159 CGATGACGCTCTGAGGTGAGCGCGCGCAGCATCATTTGAGATGATCCCAAGCAAGATGTCGA 2218
755 uAspLeuLeuGlyLysIleGln 762
2219 TGATGAATATGCGTGTCCAG 2240
RESULT 44
US-09-535-008-64
; Sequence 64, Application US/09535008
; Patent No. 6465629


```
QY 456 uAlaLeuAspLeuLeuGluArgLeuLysAsp-----Val** 469
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1496 ATACCTCAATAGCATCTCCAGCATGCCAAGGATTCAAGGAATATCAGATCCGCTCAC 1555
QY 469 *SerAspLysValLysLeuVal**AspLeuAlaPheLeuAlaProIle----- 486
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1556 AGGCAAAATCCAGAGCTGACCAAGGCGATGCCAGTCCATTCACCAACAGCGGCGGA 1615
QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
Db ::|||:||||:||||:||||:||||:||||:||||:||||:||||:
1616 GCAGAAGAAAGAGACGAGCGGATCGAGAGCGCATGCGAGGCTCATGGCTGAAGA 1675
QY 504 pGluLeuGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1676 TGAGGAGGGGTACCGCAAGCTCATC----- 1700
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1701 -----GACCAGAAAGAGGACAGCGCTGCGCTTACCTTTG-----CAGCAGACAGACA 1750
QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564
Db ::|||:||||:||||:||||:||||:||||:||||:||||:||||:
1751 GTACGTGCTAACCTCAGGAGCTGTCGGGCGAGCACAAAGCTGCCAGTCCCAAGA 1810
QY 564 aLysGlyLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1811 GAAAAGAGAGAAAAAGAAAGAGAGCGAGAAATGCAGAA---GGACAGACGCTGTC 1867
QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1868 CATTGGGCGGATGGC----- 1897
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGln 618
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1898 GACCAGCCAGATGACGACCTCCCGCTGAAGTG-----ATCCAGCTGGAGAGTG 1948
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGl 638
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
1949 GAAGATCTTCACAGGC----- 1964
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658
Db ::|||:||||:||||:||||:||||:||||:||||:||||:||||:
1965 -----ACATGCCCCCAAAAGCGCGGCGAGCTGGAGGCTCGAGATGAACCCGG 2017
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
2018 GTAT-----GAAGTAGCTCCGAGGTCTGAT----- 2042
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGl 698
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
2043 -----AGTGA 2047
QY 698 uGluLysProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSe 718
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
2048 AGAAAGTGGCTCAGAGAAGAGAGAGAGAGAG-----CAGGTAAGAGACGCCGCGAGC 2101
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGl 738
Db ::|||:||||:||||:||||:||||:||||:||||:||||:||||:
2102 AGCAGCTCTCCACC-----CTGCCGTGGAGAGAGAGAGATTCAGATCCAGACAG 2158
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaGl 755
Db ::|||:||||:||||:||||:||||:||||:||||:||||:||||:
2159 CGATGACCTCTCTGAGGTGGACGCGGCACATCATTCGATGCGCAAGCAGATGTGCA 2218
QY 755 uAspLeuLeuGlyLysIleGln 762
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
2219 TGATGAATATGGCGTGTCCACG 2240
```

RESULT 45

US-09-535-008-76

; Sequence 76, Application US/09535008

; Patent No. 6465629

```
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 2318-259
; CURRENT APPLICATION NUMBER: US/09/535,008
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 5573
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (75)..(5117)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (1780)
; OTHER INFORMATION: GenBank Accession No. 6465629 U29175 shows a C at this
; OTHER INFORMATION: position (position 1784 in GenBank) rather than
; OTHER INFORMATION: the G shown here.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (26)
; OTHER INFORMATION: Polymorphism of either T or C in this noncoding
; OTHER INFORMATION: region.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1583)
; OTHER INFORMATION: Polymorphism of A or G resulting in a silent
; OTHER INFORMATION: mutation.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1598)
; OTHER INFORMATION: Polymorphism of T or C resulting in a silent
; OTHER INFORMATION: mutation.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1892)
; OTHER INFORMATION: Polymorphism of A or G resulting in a silent
; OTHER INFORMATION: mutation.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4245)..(4349)
; OTHER INFORMATION: Addition of 105 basepairs compared to SEQ ID NO:1.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4603)..(4604)
; OTHER INFORMATION: Deletion of CAG between these two basepairs as
; OTHER INFORMATION: compared to SEQ ID NO:1 (deletion of basepairs
; OTHER INFORMATION: 4499-4501 of SEQ ID NO:1).
US-09-535-008-76
```

```
Alignment Scores:
Pred. No.: 0.000914 Length: 5573
Score: 150.50 Matches: 146
Percent Similarity: 30.35% Conservative: 81
Best Local Similarity: 19.52% Mismatches: 241
Query Match: 3.74% Indels: 280
DB: 4 Gaps: 31
```

US-09-765-272A-66 (1-763) x US-09-535-008-76 (1-5573)

```
QY 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
353 GAAAGGAATGGGGATGCGGTCAAGGGGCCCATGTGTGGATGGGGCCCGCCAGCCCAT 412
QY 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAl 207
```

413 GACACGACACTCCCAAGTTACCCC---TGCCCCCTGGGTGGCTCTGAGCATGCTCTAG 469
207 agluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsnAl 227
470 TCCAGTT-----CCAGCAGTGGCCGCTCTTCGGGGCCCGCAGATGCTCTTC 514
227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240
515 CGGGCCAGGAGTGGCCCGCTGGATGGTGTGACCCCCCAGGCGCTTGGGGCAGCAACCG 574
240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 260
575 GGGCCCCAACCCCA---TTTAACCAAGAACCCAG-----CTGCACCA 610
260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280
611 GCTCAGAGCTCAGATCATGCGCTTACAGATGCTG----- 644
280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
645 -----CCAGGGGGCAGCCCTCCCGACACCTGCAGATGGC 682
300 sPheIleProTyrGluGlnMetSerGluLeuGlyLeuArgIle----- 317
683 GGTGACGGGCAAGCGCGATGCGCGGATGCAGCAGCAGATGCCAACGCTACCTCCACC 742
317 ----- 317
743 CTCGGTGTCCGCAACAGACCGCGCTGCTGGCCCTGCGCCCGCGCGGTCCCGG 802
318 -----IleProLeuArgTyrArgSerAsnHis----- 326
803 CCGGCACCTCCAAATACAGCAGGCTCATGTATGGAGGGCCCAACATGCCTCCCCC 862
327 -----TrpValProAspSerArgProGluGlnPro----- 336
863 AGGACCTCGGGGTGCCCGCGGATGCCAGCGCAGCCTCTTGAGGGCTCCCAAGCC 922
337 -----SerProGlnSerThrProGlu----- 343
923 CTGGCTGAAGCACCCTCGCGAATGCTGTGCCCCCAGCAGCACCCTCAGAACTCAT 982
344 -----ProSerProSerProGlnProAlaProAsnPro----- 354
983 TCCCCGCGCCAAAGCGCGCTTCCCCCGCGCCCTCTGCGCTCCACCGCGCGCTC 1042
355 -----GlnProAlaProSerAs 360
1043 GCGCGTGATGCCACGACAGCCAGTCCCGCGGCGAGCGCGCCCGCCCATGGT 1102
360 nProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
1103 GCGACTGCACAGAGCAGCGCGCATCACCCTCATCCAGAACCGCGGGGCTCGACCC 1162
378 -----TyrValPheGluGluAsnGlyValSerArgTyrI1 389
1163 TGTGGAGATCTCGAGAGCGCGAGTACAGGCTGCGAGCTCGCATCGCACACCGAATTCA 1222
389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405
1223 GGAACCTGAAAACCTTCCCGGGTCCCTGGCGGGGATTTGCGAACCAAGCGCACTGA 1282
405 ----- 405
1283 GCTCAAGGCGCTCAGGCTGCTGAACCTCCAGAGCAGCTCGCGCAGAGGTGGTGTG 1342
406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423
1343 CATCGGAGGGACACAGCGCTGAGACAGCGCTCAATGTCTAAAGCCCTACCAAGCGCACAA 1402
423 uProSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAlaArgIleHisG 443

Db 1403 GCGCCAGTCCCTGCGCGAG-----GCCGCGATCACTGA 1435
QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456
Db 1436 GAAGCTGGAGAGCAGCAGAGATCGAGCAGAGCGCAAGCGCGCAGAGCAGCAGGA 1495
QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
Db 1496 ATACCTCAATGACTCTCCAGCATGCCAAGGATTTCAAGGAATATCAAGATCCGCTCAC 1555
QY 469 *SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle----- 486
Db 1556 AGGCAAAATCCAGAGCTGACCAAGCAGTGGCCACCTACCTGCCAACACGAGCGGGA 1615
QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
Db 1616 GCAGAAAGAGAGAACGAGCGGATCGAAGAGCGCATCGGAGGCTCATGGCTGAAGA 1675
QY 504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524
Db 1676 TGAGGAGGGTACCGCAAGCTCATC----- 1700
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544
Db 1701 -----GACCAGAAAGAGCAACGCGCTGCTACCTCTTG-----CAGCAGACAGACGA 1750
QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564
Db 1751 GTACGTGGCTAACCTCAGGAGCTGGTGGCGCAGCACAAAGGCTGCCAGGTCCCAAGGA 1810
QY 564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
Db 1811 GMAAAGAGAGAAAGAAAGAAAGAGCAGCAAAATGCAGAA---GGACAGACGCTGC 1867
QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaLysLysValProLeuAspAr 604
Db 1868 CATTTGGCGGGATGTC-----GAGCCTCTAGACGA 1897
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnG 618
Db 1898 GACCAGCCAGATGACGACCTCCCGGTGAAGTG-----ATCCAGTGGAGAGTGG 1948
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspG 638
Db 1949 GAAGATCTCACAGC----- 1964
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658
Db 1965 -----ACAGATGCCCAAGCGCGCAGCTGGAGGCTGCTCGAGATGAACCGGG 2017
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
Db 2018 GTAT-----GAAGTAGTCTCCGAGGTCTCAT----- 2042
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG 698
Db 2043 -----AGTGA 2047
QY 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
Db 2048 AGAAGTGGCTCAGAGAAAGAGAGAGAGAGAGAG-----GAGGAAGACAGCGCGCAGGC 2101
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG 738
Db 2102 AGCAGCGCTCCACC-----CTGCGCGTGGAGAGAGAGAGATTTCCAGATCCAGACAG 2158
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaG 755
Db 2159 CGATGACGCTCTCTGAGGTGGACCGCGCGCACATCATTTGGAATGCGCAAGATGTCGA 2218
QY 755 uAspLeuLeuGlyLysIleGln 762
Db 2219 TGATGAATATGCGGTGCTCCAG 2240

RESULT 46
 US-09-535-008-72
 ; Sequence 72, Application US/09535008
 ; Patent No. 6465629
 ; GENERAL INFORMATION:
 ; APPLICANT: Wong, Alexander K.C.
 ; APPLICANT: Tavtigian, Sean V.
 ; APPLICANT: Teng, David H.-F.
 ; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
 ; TITLE OF INVENTION: AND OTHER CANCER TYPES
 ; FILE REFERENCE: 2318-259
 ; CURRENT APPLICATION NUMBER: US/09/535,008
 ; CURRENT FILING DATE: 2000-03-23
 ; EARLIER FILING DATE: 1999-03-23
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 72
 ; LENGTH: 5576
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (75)..(5120)
 ; FEATURE:
 ; NAME/KEY: conflict
 ; LOCATION: (1780)
 ; OTHER INFORMATION: GenBank Accession No. 6465629 U29175 shows a C at this
 ; position (position 1784 in GenBank) rather than
 ; the G shown here.
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: (26)
 ; OTHER INFORMATION: Polymorphism of either T or C in this noncoding
 ; region.
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: (1583)
 ; OTHER INFORMATION: Polymorphism of A or G resulting in a silent
 ; mutation.
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: (1598)
 ; OTHER INFORMATION: Polymorphism of T or C resulting in a silent
 ; mutation.
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: (1598)
 ; OTHER INFORMATION: Polymorphism of T or C resulting in a silent
 ; mutation.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4245)..(4349)
 ; OTHER INFORMATION: This is a 105 basepair insertion as compared to
 ; OTHER INFORMATION: SEQ ID NO:1.
 US-09-535-008-72

Alignment Scores:
 Pred. No.: 0.000915 Length: 5576
 Score: 150.50 Matches: 146
 Percent Similarity: 30.35% Conservative: 81
 Best Local Similarity: 19.52% Mismatches: 241
 Query Match: 3.74% Indels: 280
 DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-72 (1-5576)
 QY 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
 Db 353 GAAGAGTGGGATGGCGTTCAGGGGGCGCATGCTGGGATGGGGCCCCCGCCAGCCCAT 412
 QY 198 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207

Db	413	GGACGAGCAGCTCCCAAGGTTACCC	TCGGCCCTGGGTGGCTGAGCATCCTCTAG	469
QY	207	agluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl		227
Db	470	TCCAGTT	CCAGCCAGTGGCGCTCTCGGGGCCCCAGATGCTTC	514
QY	227	aasnPro	AladnProArgLeuSerGluAsnHisnLe	240
Db	515	CGGGCCAGGAGTGCCCGCTGGTGTGACCCAGGCTTGGGGCAGCAGAACCG		574
QY	240	uThrValThrProThrTyrHisglnAsnGlnGlyGluAsnLysSerSerLeuLeuArgL		260
Db	575	GGGCCCAACCCCA	TTTAACCAAGAACAG	610
QY	260	uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuLeuPheAspPr		280
Db	611	GCTCAGAGCTCAGATCATGGCTACAGATCTG		644
QY	280	oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi		300
Db	645		GCCAGGGGCGAGCCCTCCCGCAGATGCCACCTGCAGATGGC	682
QY	300	sPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIle		317
Db	683	GCTGAGGCAAGCGCGCATGCCGGGATGCAGCAGCAGATGCCACGCTACTCTCCACC		742
QY	317			317
Db	743	CTCGGTGTCGCAACAGGACCGGCTGGCCCTGGCCCTGGCCCGCCCGGGTCCCGG		802
QY	318	----IleProLeuArgTyrArgSerAsnHis		326
Db	803	CCCGGCACCTCCAAATTACAGCAGGCTCATGTATGGAGGGCCCAACATGCTCTCCCC		862
QY	327	-----TrpValProAspSerArgProGluGlnPro		336
Db	863	AGGACCTCGGGCGTGCCTCCCGGATGCCAGCCAGCTCTGGAGGGGCTCCCAAGCC		922
QY	337	-----SerProGlnSerThrProGlu		343
Db	923	CTGGCTGAAGACCCATGGCGATCTGCTGCTGCCCCACAGCAGCACCCTCAGAGTGTAT		982
QY	344	----ProSerProSerProGlnProAlaProAsnPro		354
Db	983	TCCCGCGAGCAACCGGCGGCTTCCCGCGGCTCCCGTCCACCCGCGGCTC		1042
QY	355		GlnProAlaProSerAs	360
Db	1043	GCCCGTGTATGCCACCCAGACCCAGTCCCGCGGCGAGCCCGCCAGCCCGCCCTGGT		1102
QY	360	nProIleAspGlyLysLeuValLysGluAlaValArgLysValGlyAspGly		377
Db	1103	GCCATGTACCAGACAGCAGCGCATCCCCCATCCAGAACCGCGGGCTCGACCC		1162
QY	378	-----TyrValPheGluGluAsnGlyValSerArgTyrI		389
Db	1163	TGTGGAGATCTTCAGGAGCGGAGTACAGGCTCGCATCGCACACCGCAATTCA		1222
QY	389	eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys		405
Db	1223	GGAACTTGAAACCTTCCCGGGTCCCTGGCGGGGATTTGCGAACCAAGCGACCATTGA		1282
QY	405			405
Db	1283	GCTCAAGGCCCTCAGGCTGCTGAACCTTCAGAGCGAGCTCGCCGAGGAGGTGGTGTG		1342
QY	406	-LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys		423
Db	1343	CATCGCGGGACACAGCGCTGGAGACAGCCCTCAATGCTAAGGCTTACAGCGACGAA		1402
QY	423	uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisG		443

Db 1403 GCCCGAGTCCCTCGCGAG-----GCCCGCATCACTGA 1435
Qy 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValaAspPheG1 456
Db 1436 GAAGCTGGAGAACGACAGAGATCGAGCAGGACGCGCGGCGAGACACACGGA 1495
Qy 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
Db 1496 ATACCTCAATAGATCTTCAGCATGCCAGGATTCAGGAATATCACAGATCCGTCAC 1555
Qy 469 *SerAspLysValLeuVal***AspIleLeuAlaPheLeuAlaProIle----- 486
Db 1556 AGGCAAAATCCAGAGCTACCAAGGAGTGGCCAGCTACCATCGCAACACGCGGCGGA 1615
Qy 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyThrAspAs 504
Db 1616 GCAGAAGAAGAGAACGAGCGGATCGAAGAGAGCGCATCGGAGGCTCATGGCTGAAGA 1675
Qy 504 pGluIleGlnValAlaLysLeuAlaGlyLysTyThrThrGluAspGlyTyrllePheAs 524
Db 1676 TGAGGAGGGTCCGCGAGTCTAC----- 1700
Qy 524 pProArgAspIleThrSerAspGluGlyAspAlaTyValThrProHisMetThrHisSe 544
Db 1701 -----GACCAGAGAGAGCAAGCGCTCGCTACCTCTTG---CAGCAGACAGACGA 1750
Qy 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyAl 564
Db 1751 GTACGTGGCTAACTCAGCGAGCTGTGGCGGACGACAGAGCTGCCAGGTGCGCAAGGA 1810
Qy 564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
Db 1811 GAAAG 1867
Qy 584 aLysGlyAlaGluAlaIleTyAsnArgValLysAlaAlaLysLysValProLeuAspAr 604
Db 1868 CATTGGCGCGATGGC-----GAGCCTCTAGACGA 1897
Qy 604 g-----MetProTyrAsnLeuGlnTyThrValGluValLysAsnG1 618
Db 1898 GACCAGCCAGATGAGCGACCTCCCGGTGAAGGTG-----ATCCACGTGGAGAGTGG 1948
Qy 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspG1 638
Db 1949 GAAGATCCCTCAGACG----- 1964
Qy 638 uGlyLeuTyGluAlaProLysGlyTyThrLeuGluAspLeuLeuAlaThrValLysTy 658
Db 1965 -----ACAGATGCCCCAAAGCGCGGACCTGGAGGCTGGCTCGAGATGAACCCGGG 2017
Qy 658 rTyValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
Db 2018 GTAT-----GAAGTAGTCCGAGGTCTGTAT----- 2042
Qy 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG1 698
Db 2043 -----AGTGA 2047
Qy 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
Db 2048 AGAAGTGGCTCAG 2101
Qy 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG1 738
Db 2102 AGCAGACCTCCCAACC---CTGCCCGTGGAGAGAGAGAGAGAGATTCAGATCCAGACAG 2158
Qy 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaG1 755
Db 2159 CGATGAGTCTCTGAGGTGGAGCGCGGCACATCATTTGAGATGCCAAGAGATGTGGA 2218
Qy 755 uAspLeuLeuGlyLysIleGln 762
Db 2219 TGATGAATATGGCGTGTCCCGAG 2240

RESULT 47

US-08-446-137B-1
; Sequence 1, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/446,137B
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-May-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMaisters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Peptostreptococcus asaccharolyticus
; STRAIN: 1018
; FEATURE:
; NAME/KEY: exon
; LOCATION: 103..3186
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start= 280
; OTHER INFORMATION: /product= "mature protein L"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 103..3186
; OTHER INFORMATION: /codon_start= 103
; OTHER INFORMATION: /product= "Immature protein L"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 280..3183
; OTHER INFORMATION: /codon_start= 280
; OTHER INFORMATION: /product= "mature protein L"
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 208..279
US-08-446-137B-1

2Y 545 HistTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAla 564
Db 2581 GAATGGTTATTAAAGTAATGCT-----AAAGAAGCTGCAATCAAGAATTA 2625
2Y 565 LysGluLysGluLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAla 584
Db 2626 AAAGAAGCGAGGAATCACT----- 2643
2Y 585 LysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArg 604
Db 2644 -----GCTGAATATCTATTCACTTAATCAACAAGCAAAACAGTAGAA----- 2688
2Y 605 MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHis 624
Db 2689 -----GGCTGAGATCATTTAAAGAACCAAGAAATCTTTAAAGACAC 2727
2Y 625 TyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAlaPro 644
Db 2728 GCTGAATAAACCGCGGAAACCCAGGAATCACAAATTGATGATGGTTATTAAAGAACGCT 2787
2Y 645 LysGlyTyrThrLeuGluAspLeuLeu-----AlaThrValLysTyrTyrValGlu 661
Db 2788 AAAGAAGATGCAATTAAAGATTAAGAAAGACAGGAATTACTCTGACATATACTTTGAT 2847
2Y 662 HisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGln 681
Db 2848 GCTATCAACAAA-----GCAAAAACCTATTGAAGCGGTAGAA 2883
2Y 682 ArgAsnLysAsn-----GlyGlnAlaAspThrAsnGlnThrGluLysPro 696
Db 2884 GCATTAAAGATGAATCTTTAAAGGCTCATATAAAGATGAAGAACACAGGTATAAACA 2943
2Y 697 SerGluGlu-----LysProGlnThrGluLysPro-----GluGluGluThrProArgGluGlu 714
Db 2944 GGTGAAGCAAAACCAAGCAAGATATAAACCAGGTGAAGATATAAACCAGAAACAAA 3003
2Y 715 LysPro-----GlnSerGluLysProGluSerProLysProThrGluGluProGluLysSer 733
Db 3004 AAACCTGGTGAAGATAAATAAACCAGAGACAAACACAGGTAAACAGATTAAGATTCT 3063
2Y 734 ProGluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGlu 753
Db 3064 CCAATAAAGAAAGAAAAA-----GCTAAATATCAAAAGAGCTGGTAGCGAA 3108
2Y 754 AlaGluAspLeu 757
Db 3109 GCTGAATCTTTA 3120

RESULT 48
JS-09-535-008-62
Sequence 62, Application US/09535008
Patent No. 6465629
GENERAL INFORMATION:
APPLICANT: Wong, Alexander K.C.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.-F.
TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
TITLE OF INVENTION: AND OTHER CANCER TYPES
FILE REFERENCE: 2318-259
CURRENT APPLICATION NUMBER: US/09/535,008
CURRENT FILING DATE: 2000-03-23
EARLIER APPLICATION NUMBER: U.S. 60/125,806
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 62
LENGTH: 5471
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (75)...(1703)
JS-09-535-008-62

Alignment Scores:
Pred. NO.: 0.00111 Length: 5471
Score: 149.50 Matches: 146
Percent Similarity: 30.35% Conservative: 81
Best Local Similarity: 19.52% Mismatches: 241
Query Match: 3.72% Indels: 280
DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-62 (1-5471)

QY 177 GluAspThrGlyAspAlaTyr-----lIleValProHis 187
Db 353 GAAAGGAATGGGATGCGGTCAAGGGGGCCATCTGGGATGGGGCCCCGCCCGCAGCCCCAT 412
QY 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207
Db 413 GGACGACGACTCCCAAGGTATACCCC---TCGCCCTGGTGGCTCTGACATGCTCTAG 469
QY 207 agluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl 227
Db 470 TCCAGTT-----CCAGCCAGTGGCGGCTCTCGGGGCCCCAGATGCTTC 514
QY 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240
Db 515 CGGGCCAGAGGTGCCCGCTGGATGGTCTGACCCCGAGGCTTGGGGCAGCAGAACCG 574
QY 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 260
Db 575 GGGCCCAACCCCA---TTTACCAGAACCG-----CTGCACCA 610
QY 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280
Db 611 GCTCAGAGCTCAGATCATGGCTACAGATGCTG----- 644
QY 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
Db 645 -----GCCAGGGGGCAGCCCTCCCGCAGCACCTGCAGATGGC 682
QY 300 sPheIleProTyrGluGlnMetSerGluLeuLysArgIleAlaArgIle----- 317
Db 683 GGTGAGGGCAAGCGCGGATGCCCGGATGAGCAGCAGAGTGCACACGCTACCTCCACC 742
QY 317 ----- 317
Db 743 CTCGGTGTCCGCAACAGGACCGCGCTGGCCCTGGCCCTGGCCCGCGCGGTCCCGG 802
QY 318 ----IleProLeuArgTyrArgSerAsnHis----- 326
Db 803 CCGGSCACCTCCAAATTACAGCAGGCTCATGTATGGAGGGGCCCAACATGCTCCTCCCC 862
QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336
Db 863 AGGACCTCGGGCGTGGCCCGGGATGCCAGGCCAGCTCTCTGGAGGGCCTCCCAAGCC 922
QY 337 -----SerProGlnSerThrProGlu----- 343
Db 923 CTGGCTGAAGACCCATGGCGGAATGTGCTGCCCCACAGCACCCTTCAGAGTGTAT 982
QY 344 ----ProSerProSerProGlnProAlaProAsnPro----- 354
Db 983 TCCCCCGCAGCAACGGCGCGGCTTCCCGCGCCCTCCCGCGCTCCCGCGCGCTC 1042
QY 355 -----GlnProAlaProSerAs 360
Db 1043 GCCCGTGTATGCCACCGCAGACCCAGTCCCGCGGCGAGCCCGCGCGCGCTCATGTT 1102
QY 360 nProfileaspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
Db 1103 GCCACTGCACCAAGACGCGGATCACCCCGCATCCAGAAAGCCCGGGGCTCCAGCC 1162
QY 378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389

Db 1163 TGTGGAGATCTCTGAGGAGCGGAGTACAGGCTGCGAGGCTCGCATCGCACCCGAATTCA 1222
QY 389 eProAlaIysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerIlys----- 405
Db 1223 GGAACCTTGAACACCTTCCCGGGTCCCTGCGGGGATTTGGAAACCAAGGACCATTTGA 1282
QY 405 ----- 405
Db 1283 GCTCAAGCCCTCAGGCTGTGAATCTCCAGAGGCGAGTGGCCAGGAGGTGGTGGTGTG 1342
QY 406 -LeuAlaIysGlnGluSerLeuSerHisIysLeuGlyAlaIys-----LysThrAspLe 423
Db 1343 CATCGGAGGACACAGCGCTGGAGACAGCCCTCAATCTTAAGGCTACAGCGGAGCA 1402
QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisG 443
Db 1403 GCGCCAGTCCCTGCGCGAG-----GCCCGCATCACTGA 1435
QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456
Db 1436 GAAGCTGGAGAGCAGCAGAGATCGAGCAGGCGCAAGCGCCGCGCAGAACCCAGGA 1495
QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val** 469
Db 1496 ATACCTCAATAGCATTTCCAGCATGCCAAGATTCAGGAATATCACAGATCGTCTAC 1555
QY 469 *SerAspLysValIysLeuVal***AspIleLeuAlaPheLeuAlaProIle----- 486
Db 1556 AGGCAAAATCCAGAAAGTGCACCAAGGCGAGTGCACATGCCAACACAGGCGGGA 1615
QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
Db 1616 GCAGAGAAAGAGAACGAGCGGATTCGAGAGGAGCGCATGCGGAGGCTCATGCTGA 1675
QY 504 pGluIleGlnValAlaIysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524
Db 1676 TGAGGAGGATACCGCAAGCTCATC----- 1700
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisE 544
Db 1701 -----GACTAGAAGAGGACAGACGCGCTGGCTACTCTCTG---CAGCAGACAGAGA 1750
QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564
Db 1751 GTACGTGGCTTAACCTCACGAGGCTGTGTGCGCAGCAGACAGGCTGCCAGGTCGCAAG 1810
QY 564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
Db 1811 GAAAAGAGAAAGAAAGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1867
QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604
Db 1868 CATGGGCGGATGGC-----GAGCCTCTAGAGA 1897
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnG 618
Db 1898 GACCAGCCAGATGAGCAGCCTCCCGGTGAAGGTG-----ATCCAGCTGAGAGGTG 1948
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspG 638
Db 1949 GAAGATCCTCACAGGC----- 1964
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658
Db 1965 -----ACAGATGCCCCAAAGCCGGCAGCTGGAGCCTGGCTCAGATGAACCCGGG 2017
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
Db 2018 GTAT-----GAAGTAGTCCCGAGGTCTGTAT----- 2042
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG 698
Db 2043 -----AGTGA 2047

QY 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
Db 2048 AGAAAGTGTCTCAGAAAGAGAGAGAGAGAGAGAG-----GAGGAAGAGCAGCGCAGGC 2101
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG 738
Db 2102 AGCAGACGCTCCACCC-----CTGCCCTGGAGGAGAGAGAGATTCAGCATCCAGACAG 2158
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaG 755
Db 2159 CGATGACGCTCTGAGGTGAGCGCGGCGCACATCATTTGAGATGCCAAGCAAGATGTGCA 2218
QY 755 uAspLeuLeuGlyLysIleGln 762
Db 2219 TGATGAATATGCGGTGTCCAG 2240
RESULT 49
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: TUBERCULOSIS
; FILE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 48.7 Length: 4403765
Score: 149.50 Matches: 161
Percent Similarity: 30.71% Conservative: 89
Best Local Similarity: 19.78% Mismatches: 284
Query Match: 3.72% Indels: 282
DB: 3 Gaps: 37
US-09-765-272A-66 (1-763) x US-09-103-840A-2 (1-4403765)
QY 129 ArgThrLysGluGluIleLysArgGlnLysGlnGluArgSerHisAsnHisAsnSer--- 147
Db 2479156 AGGCCAGAGGTGGACCTGCGCGCAGCAGATCGAGGAGTCAACAGAAATGCGCTT 2479215
QY 148 ---ArgAlaAspAsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrThrAspAsp 166
Db 2479216 CTCGCTCCAGATGCGGCGCATCGGTGAGAGCGTCCACCGAGG---GAC 2479260
QY 167 GlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValPro 186
Db 2479261 GGTAC-----CCG 2479269
QY 187 HisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAla 206
Db 2479270 CT-GGCTCAACAGAGGAGCGACACGCTCGAAGTCTCGACGAGCCCTCG----- 2479316
QY 207 AlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsn 226
Db 2479317 ---TGAGG-----TGTCGACCGCAAGGTCGACACCG----- 2479346

2Y 227 AlaAsnProAlaGlnProArgLeuSerGluAsnHisLenLeuThrValThrProThrTyr 246
2Y 2479347 AATCCCTTCGCGCGCGGTGTGTGACCAAGATCATCGCCAGGAGATGACACGG 2479406
2Y 247 HisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu 266
2Y 2479407 TCGAGGTGGCGCGAGCTCGCTGTCTATGGCGAGCCCAAGATGCGCGGAGGCGG 2479463
2Y 267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286
2Y 2479464 CGGCCCGGACCCGAGAAAGTCC-----CTGCGGCCCAACCCGAGTCCAAGCGG 2479514
2Y 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
2Y 2479515 -----CACCCGAACACCAC-----CGG 2479532
2Y 307 MetSerGluLeuGluLysArgIleAlaArgIlePro----- 319
2Y 2479533 TCCACCGAGCTCGGAGCGCTCTGTGTGGCGATGCCAAGCCGGTGTGATGCCCGAGC 2479592
2Y 320 -----LeuArg 321
2Y 2479593 TCGGCGAATCGGTGACCGAGGGACCGTTCATTCGTGCTGAAGAAGATCGGGATTCGG 2479652
2Y 322 TyrArg-----SerAsnHis-----TrpValProAspSerArg 332
2Y 2479653 TTCAGGTTGACGAGCCACTCGTGAGGTGTCCACCGACAAGGTGGACACCGAGATCCCGT 2479712
2Y 333 Pro----- 333
2Y 2479713 CCCCGTGGCTGGGTCTTGGTCAGTATACAGCGCCGAGGAGCGCCACCGTGCCTGCG 2479772
2Y 334 -----GluGlnProSerProGlnSerThr 341
2Y 2479773 GCGCGAGTTGGCCCGATCGGTGTCTGCTGCCGACATCGCGCGCGCCCGCCAGC 2479832
2Y 342 ProGluProSerProSerProGlnProAlaProAsnPro-----GlnProAlaProSer 359
2Y 2479833 CGGCACCCAAAGCCGTCCTCCCGAG-CCAGCGCCGAGCGCGGAGGCGCAACCGCACCATCG 2479891
2Y 360 AsnProIleAspGlu-----LysLeuValLysGluAlaVal 371
2Y 2479892 CGCGCGCGCGCCACGACGCGGTGCGCGGAGCGCGCACCGTACGTGACGCGCGCTGTG 2479951
2Y 372 ArgLysValGlyAspGlyTyrValPheGluAsnGlyValSerArgTyrIleProAla 391
2Y 2479952 CGAAAGCTGGG-----TCGAAACACAAATC----- 2479978
2Y 392 LysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSer 411
2Y 2479979 ---GACCTCGCGGGGTGACCGGCACCGGAGTGGGTGTGTCATCCGCAACAGAGATGTG 2480035
2Y 412 LeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAsp----- 427
2Y 2480036 CTGGCGCGCGCTGAGCAAAAGAGCGGCGGAAAGCACCGGCGCGCGCGCGCC 2480095
2Y 428 -----ArgGlu 429
2Y 2480096 GCGCGCGCGCGCCGAAAGCGCGCTGCTCCCTGCGCGCGCGGTGGCACATCATCGGGCG 2480155
2Y 430 PheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAsnLys 449
2Y 2480156 ACCACCCAGAGGCC-----AGCGGATTCGTGAG---ATCAGCGCCCAACAG 2480200
2Y 450 GlyArg-----GlnValAspPheGlu 456
2Y 2480201 ACCCGGAATCTTTGAGGCACGCGCACAGCTGACACAAACCCATGAGGTGCGATGACC 2480260
2Y 457 AlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal 476
2Y 2480261 AAGATCGTGGGCTACGGGCGCGCGGCAAGCGCGGTTCGCCGAGCGGTGAGCGGTGAAC 2480320
2Y 477 ***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsn 496

RESULT 50

US-09-103-840A-1

Db 2480321 CTGACCTTCTCGCGGTCTTCGCCAAGCGCGTGATGATGCCCTCAAGATTCAACCCGAAC 2480380
QY 497 AlaGlnIleThrTyrThrAspAspGluIleGlnValAla-----LysLeu 511
Db 2480381 ATCAACGTAGCTACACAGGAGACCAAGGAGATCACCTACTACAGCGCGGACACCTA 2480440
QY 512 AlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProArgAspIleThrSerAsp 531
Db 2480441 GGATTCGTGTGACACCGAGCGAGGCGCTGCTC-----TCCCGGTTCATCCACGAC 2480491
QY 532 GluGlyAspAlaTyrValThrProHisMetThrHisSerHisTrpIleLysLysAspSer 551
Db 2480492 GCCGCGCATCTG-----TCA 2480506
QY 552 LeuSerGluAlaGluArgAlaAlaGln-----AlaTyrAlaLysGluLysGlyLeu 569
Db 2480507 CTGGCGGTCTCGCGCGCGGATGCGCGGATGCGCGCGGTGCGCGCAACCTG 2480566
QY 570 ThrProSerThrAspHisGlnAspSerGly-----AsnThrGluAla 584
Db 2480567 AAACCC-----GACGAGTTGTCCGGCGGACCTTCACCATCACCAATCGGTAGC 2480617
QY 585 LysGlyAla-----GluAlaIleTyrAsnArgVal 594
Db 2480618 CAGGCGCGGTGTTCGACACACCCCGATCTCTGTTCCGCGCAGCGCGCATGTGGGCACC 2480677
QY 595 LysAlaAlaLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614
Db 2480678 GGGCGCATCGTCAACCGCGCGGTGTGTCGATGCCAGCGGCAACGAGTCGATCGGG 2480737
QY 615 ValLysAsnGlySerLeuIleIleProHisTyrAspHis----- 627
Db 2480738 CTGGCTCGGTCTGTACTCTCCCTGCTGACCTATGACCATCGGTCTCATCGCGCGCGAC 2480797
QY 628 -----TyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAla 643
Db 2480798 GCGCGACGCTTCTCACCACGATCAAGACCGC---CTCGAAGAGGAGCGGTTCGAGGCC 2480854
QY 644 ProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrVal----- 650
Db 2480855 GATTAGGACTGTGATGCCAACGCGGTGTGTCGATCGCGGTTCGTCTGGCTGTATCG 2480914
QY 661 -----GluHisProAsnGluArgProHisSerAspAsnGly 672
Db 2480915 GCTCTGCCCTGACCGCGCGGTGCGCGCGCGCCACACACCG-----TGC 2480959
QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn-----GlyGln 687
Db 2480960 TCGGATCGTGGCGCGCGCACCTCGGAATTCGGAAGACTGCACTGGAATCCGAAAGCG 2481019
QY 688 AlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysProGlu 707
Db 2481020 CGGAATTCGATCCGACGCGCTCACCGATGTGCGCGGTGTGTCACCTCTCGCGGTCA 2481079
QY 708 GluGlu-----ThrProArgGlu 713
Db 2481080 ACATCGCCCGCGTGTGTCGGGGCTTTCAAACAGAGCGCTGCGCGACACCGCGATCA 2481139
QY 714 GluLysProGlnSerGluLysProLysProThrGluGluProGluGluSer 733
Db 2481140 CACCCACGAGGTGTATCCCGCGAGTGTGCGCGCGCGGTGCTGCTGATCAACG 2481199
QY 734 ProGluGlu-----SerGluGluProGln-----ValLeuThrGluLys 746
Db 2481200 CCAGCGGTGGGTACTACTACGGAACACCAAGACCGGTGTGTCAGCAAAACGACTCGG 2481259
QY 747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLys 760
Db 2481260 CGGGAACAGGTTTTCTGCCCGACGCTGTGCTGACTGGGAAA 2481301

; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 ; US-09-103-840A-1

 Alignment Scores:
 Pred. No.: 48.8
 Score: 149.50
 Percent Similarity: 30.71%
 Best Local Similarity: 19.78%
 Query Match: 3.72%
 DB: 3

US-09-765-272A-66 (1-763) x US-09-103-840A-1 (1-4411529)
 QY 129 ArgThrLysGluGluLeuLysArgGlnLysGlnGluArgSerHisAsnSer--- 147
 Db 2481911 AGCCAGAGAGTGGACCTCGCCGACCGACGATCGATCGAGGAGTCAACAGAAATGGCCTT 2481970

 QY 148 ---ArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrThrAspAsp 166
 Db 2481971 CTCCTCCAGATCCCGCACTCGGTGAGACGCTACCGAGGG---GAC 2482015

 QY 167 GlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValPro 186
 Db 2482016 GGTTC--- 2482024

 QY 187 HisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 206
 Db 2482025 CT-GGCTCAACAGGAAGGACACGCTCGAATCGACGAGCCCTCG--- 2482071

 QY 207 AlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsn 226
 Db 2482072 ---TGGAGG--- 2482101

 QY 227 AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyr 246
 Db 2482102 AATCCCTCCGCGCGGGGTGCTGCACCAAGATCATCGCCAGGAGATGACACGG 2482161

 QY 247 HisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlalysProLeu 266
 Db 2482162 TCGAGGTGCGCGGAGTCTGCTGTCATTGTCGACGCCAAGGATGCGCGGAGCG--- 2482218

 QY 267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286
 Db 2482219 CGGCCCGGACCCGAGAAGTCC---CTCGGCCCAACCCGATGTCACGCGG 2482269

 QY 287 ThrAlaArgGlyValAlaProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
 Db 2482270 ---CACCCGACACAC--- 2482287

 QY 307 MetSerGluLeuGlyLysArgIleAlaArgIlePro--- 319
 Db 2482288 TCCAAACGACGTCGCGGCGCTGCTGGTGGCGATGCCAAGCGGCTGATGCCCGGAC 2482347

 QY 320 --- 321
 Db 2482347 --- 2483372

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: May 1, 2004, 02:02:14 ; Search time 163 Seconds
(without alignments)
7796.549 Million cell updates/sec

Title: US-09-765-272A-65

Perfect score: 2290

Sequence: 1 TTGTCCTATGAACCTGTC.....TACTTGGAAAAATCCAGGAT 2290

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq*

5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2285	99.8	2290	3	US-08-961-083-65
2	2285	99.8	2290	4	US-09-536-784-65
3	2220.2	97.0	2531	4	US-09-468-656A-11
4	1999.6	87.3	2531	4	US-09-468-656A-5
5	1999.2	87.3	8195	4	US-08-961-527-94
6	1376.6	60.1	2359	4	US-08-961-527-243
7	990.6	43.3	2451	4	US-09-468-656A-9
8	987.6	43.1	2389	3	US-08-961-083-55
9	987.6	43.1	2389	4	US-09-536-784-55
10	794.8	34.7	973	4	US-08-961-527-355
11	677	29.6	1684	4	US-08-961-527-258
12	401	17.5	1342	3	US-08-961-083-181
13	401	17.5	1342	4	US-09-536-784-181
14	401	17.5	1455	4	US-09-468-656A-7
15	401	17.5	6867	4	US-08-961-527-192
16	272.4	11.9	841	2	US-08-743-637B-34
17	272.4	11.9	841	3	US-08-526-840B-34
18	64.4	2.8	7218	1	US-08-232-463-14
19	56.2	2.5	2223	1	US-08-257-073-4
20	51.4	2.2	865	4	US-09-640-419C-6
21	47.8	2.1	31096	4	US-08-956-171E-59
22	46.6	2.0	2787	4	US-09-134-000C-3185
23	45.4	2.0	5340	4	US-09-627-122-21
24	44.8	2.0	3337	1	US-08-072-610-1
25	44.8	2.0	3337	2	US-08-719-822B-1
26	44.8	2.0	3337	3	US-09-092-458-1
27	44.2	1.9	43795	3	US-08-742-185-101

101 38.2 1.7 289 3 US-09-244-796-17 Sequence 17, Appl
102 38.2 1.7 966 2 US-08-766-738-2 Sequence 2, Appl
103 38.2 1.7 966 2 US-09-262-610-2 Sequence 2, Appl
104 38.2 1.7 3006 4 US-09-762-724-3 Sequence 3, Appl
105 38.2 1.7 5394 3 US-08-688-376-1 Sequence 1, Appl
106 37.8 1.7 387 4 US-09-621-976-8572 Sequence 8572, Ap
107 37.6 1.6 2169 4 US-09-434-408-3 Sequence 3, Appl
108 37.6 1.6 2447 2 US-09-014-969-14 Sequence 14, Appl
109 37.6 1.6 3135 4 US-09-107-532A-1575 Sequence 1575, Ap
110 37.6 1.6 4322 2 US-08-537-342-1 Sequence 1, Appl
111 37.6 1.6 10640 4 US-09-417-485D-5 Sequence 5, Appl
112 37.4 1.6 6617 4 US-09-689-065B-1 Sequence 1, Appl
113 37.2 1.6 658 4 US-09-286-981B-25 Sequence 25, Appl
114 37 1.6 1203 4 US-09-134-001C-1483 Sequence 1483, Ap
115 37 1.6 1695 3 US-09-308-022-5 Sequence 5, Appl
116 37 1.6 2780 4 US-09-220-132-176 Sequence 176, Ap
117 37 1.6 2780 4 US-09-220-132-178 Sequence 178, Ap
118 36.8 1.6 4081 3 US-08-999-774A-1 Sequence 1, Appl
119 36.6 1.6 240 1 US-08-628-417-6 Sequence 6, Appl
120 36.6 1.6 657 4 US-09-527-345-3 Sequence 3, Appl
121 36.6 1.6 2167 4 US-09-216-393B-326 Sequence 326, Ap
122 36.6 1.6 2167 4 US-09-216-393B-328 Sequence 328, Ap
123 36.6 1.6 2392 4 US-09-216-393B-331 Sequence 331, Ap
124 36.6 1.6 2393 4 US-09-216-393B-329 Sequence 329, Ap
125 36.6 1.6 5562 4 US-10-204-708-63 Sequence 63, Appl
126 36.4 1.6 888 4 US-09-134-001C-1629 Sequence 1629, Ap
127 36.4 1.6 956 4 US-08-714-741-36 Sequence 36, Appl
128 36.4 1.6 1056 4 US-09-328-352-2030 Sequence 2030, Ap
129 36.4 1.6 1990 4 US-08-714-741-35 Sequence 35, Appl
130 36.4 1.6 2085 1 US-08-072-070-1 Sequence 1, Appl
131 36.4 1.6 2085 1 US-08-465-746-1 Sequence 1, Appl
132 36.4 1.6 2085 1 US-08-214-164-1 Sequence 1, Appl
133 36.4 1.6 2085 1 US-08-469-434-1 Sequence 1, Appl
134 36.4 1.6 2085 1 US-08-214-222-1 Sequence 1, Appl
135 36.4 1.6 2085 2 US-08-467-852A-1 Sequence 1, Appl
136 36.4 1.6 2085 2 US-08-468-718-1 Sequence 1, Appl
137 36.4 1.6 2085 2 US-08-246-636-1 Sequence 1, Appl
138 36.4 1.6 2085 2 US-08-247-491A-1 Sequence 1, Appl
139 36.4 1.6 2085 2 US-08-319-795-1 Sequence 1, Appl
140 36.4 1.6 2085 2 US-08-468-985-1 Sequence 1, Appl
141 36.4 1.6 2086 3 US-08-312-949-1 Sequence 1, Appl
142 36.4 1.6 2086 3 US-08-446-201-2 Sequence 2, Appl
143 36.4 1.6 4766 5 PCT-US93-07261-10 Sequence 10, Appl
144 36.2 1.6 429 3 US-08-847-065-20 Sequence 20, Appl
145 36.2 1.6 1071 4 US-09-134-000C-3347 Sequence 3347, Ap
146 36.2 1.6 10993 4 US-08-961-527-15 Sequence 15, Appl
147 36.2 1.6 72604 4 US-09-268-992-7 Sequence 7, Appl
148 36.2 1.6 72604 4 US-09-657-474-7 Sequence 7, Appl
149 36 1.6 1436 4 US-09-023-655-873 Sequence 873, Ap
150 36 1.6 1453 4 US-08-714-741-33 Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-08-961-083-65
; Sequence 65, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: Hp Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2290 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-65

Query Match 99.8%; Score 2285; DB 3; Length 2290;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTCTTCTATGAAGTGGTGGTCTCACCAGCTGGTTCAGGTTAAGGAAGTCTTAATCGAGT 60
Db 1 TTCTTCTATGAAGTGGTGGTCTCACCAGCTGGTTCAGGTTAAGGAAGTCTTAATCGAGT 60

QY 61 TTCTTATATAGATGGTGGTATCAGGCTGGTCAAAAGGCGAGAAACTTGACACAGATGAAGT 120
Db 61 TTCTTATATAGATGGTGGTATCAGGCTGGTCAAAAGGCGAGAAACTTGACACAGATGAAGT 120

QY 121 CAGTAAGAGGAGGGGATCAACCGCAAAATGTTNATCAAGATTACGGATCAAGGTTA 180
Db 121 CAGTAAGAGGAGGGGATCAACCGCAAAATGTTNATCAAGATTACGGATCAAGGTTA 180

QY 181 TGTGACCTCTCATGGAGACCATTTACTATTACTAATATGGCAAGGTTCTTATGATGCCAT 240
Db 181 TGTGACCTCTCATGGAGACCATTTACTATTACTAATATGGCAAGGTTCTTATGATGCCAT 240

QY 241 CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTCAGTTTGAAGGATTCAGACATTTGT 300
Db 241 CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTCAGTTTGAAGGATTCAGACATTTGT 300

QY 301 CAATGAAATCAAGGGTGGTGGTATGTCTATTAAAGGTTAAACGGTAAATCTATGTTACCTTAA 360
Db 301 CAATGAAATCAAGGGTGGTGGTATGTCTATTAAAGGTTAAACGGTAAATCTATGTTACCTTAA 360

QY 361 GGATGAGCTCATGGGATTAATATTCGGACAAAGAGAGATTAAACGTCAAGACAGA 420
Db 361 GGATGAGCTCATGGGATTAATATTCGGACAAAGAGAGATTAAACGTCAAGACAGA 420

QY 421 ACCGAGTCAATAATCATAACTCAAGAGCAGATAATGCTGTTGCTCAGCCAGGCCAAGG 480
Db 421 ACCGAGTCAATAATCATAACTCAAGAGCAGATAATGCTGTTGCTCAGCCAGGCCAAGG 480

QY 481 ACCTTATCAACCGGATGATGGGTATATCTTCAATGCAATCTGATATCATTTGAGGACACGG 540
Db 481 ACCTTATCAACCGGATGATGGGTATATCTTCAATGCAATCTGATATCATTTGAGGACACGG 540

QY 541 TGATGCTTATATCGTTCCTCAACGGGACCATTAACATTAATTCCTTAAGATGAGTTATC 600
Db 541 TGATGCTTATATCGTTCCTCAACGGGACCATTAACATTAATTCCTTAAGATGAGTTATC 600

QY 601 AGCTAGGAGTTAGCTGCTCGAGAGCTATTGGAATGGGAAGCAGGATCTCGTCCCTTC 660
Db 601 AGCTAGGAGTTAGCTGCTCGAGAGCTATTGGAATGGGAAGCAGGATCTCGTCCCTTC 660

QY 661 TTCAAGTTCTAGTTATTAATGCAATCCAGCTCAACCAAGATTGTGAGAGAACCACAATCT 720


```

; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2290 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-09-536-784-65

Query Match          99.8%; Score 2285; DB 4; Length 2290;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTCTCTATGAACCTTGGTCGTACCAAGCTGGTCAGGTTTAAGAAAGAGTCTTAATCGAGT 60
Db      1 TTGTTCTCTATGAACCTTGGTCGTACCAAGCTGGTCAGGTTTAAGAAAGAGTCTTAATCGAGT 60

QY      61 TTCTTTATATAGATGGTGATCAGGCTGTCTAAAGCGCAGAAACCTTGACACCAGATGAAGT 120
Db      61 TTCTTTATATAGATGGTGATCAGGCTGTCTAAAGCGCAGAAACCTTGACACCAGATGAAGT 120

QY      121 CAGTAAAGAGGGAGGGATCAACGCCGAACAAATNGTNATCAAGATTACGGATCAAGGTTA 180
Db      121 CAGTAAAGAGGGAGGGATCAACGCCGAACAAATNGTNATCAAGATTACGGATCAAGGTTA 180

QY      181 TGTGACCTCTCATGAGAGACCATTAATCATTAATTAATGCGMAGGTCCTGTATGATGCCAT 240
Db      181 TGTGACCTCTCATGAGAGACCATTAATCATTAATTAATGCGAAGGTCCTGTATGATGCCAT 240

QY      241 CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCAGTTGAAGGATTACAGACATTGT 300
Db      241 CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCAGTTGAAGGATTACAGACATTGT 300

QY      301 CAATGAAATCAAGGGTGGTTATGTCTAATTAAGGTAACCGTAAATATCTATGTTTACCTTAA 360
Db      301 CAATGAAATCAAGGGTGGTTATGTCTAATTAAGGTAACCGTAAATATCTATGTTTACCTTAA 360

QY      361 GGATGACGCTCATGCGGATTAATTTCCGAACAAAGAGAGATTAAACGTGAGAGCAGGA 420
Db      361 GGATGACGCTCATGCGGATTAATTTCCGAACAAAGAGAGATTAAACGTGAGAGCAGGA 420

QY      421 ACCGAGTCTAATAATCAATAACTCAAGAGCAGATATGCTGTTGCTGCAGCAGAGCCCCAAGG 480
Db      421 ACCGAGTCTAATAATCAATAACTCAAGAGCAGATATGCTGTTGCTGCAGCAGAGCCCCAAGG 480

QY      481 ACGTTATAACCGGATGATGGGTATATCTTCAATGCACTCTGATATCATTTGAGGACAGGG 540
Db      481 ACGTTATAACCGGATGATGGGTATATCTTCAATGCACTCTGATATCATTTGAGGACAGGG 540

QY      541 TGATGCTTTATCGTTCTTCACGCGGACCATTAACATTACATTCTTAAGGATCAGGTTATC 600
Db      541 TGATGCTTTATCGTTCTTCACGCGGACCATTAACATTACATTCTTAAGGATCAGGTTATC 600

QY      601 AGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGAAGCAGGGAATCTCGTCCCTTC 660
Db      601 AGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGAAGCAGGGAATCTCGTCCCTTC 660

QY      661 TTCAAGTTCTAGTTTAAATGCAAAATCCAGCTCAAACGAAGTTGTGAGAGAACCAAAATCT 720
Db      661 TTCAAGTTCTAGTTTAAATGCAAAATCCAGCTCAAACGAAGTTGTGAGAGAACCAAAATCT 720

QY      721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTACGTGA 780
Db      721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTACGTGA 780

QY      781 ATTGTATGCTTAAACCCCTTATCAGAAACCCCATGTGGAATCTGTAGGCCCTTATTTTCGACCC 840
Db      781 ATTGTATGCTTAAACCCCTTATCAGAAACCCCATGTGGAATCTGTAGGCCCTTATTTTCGACCC 840

QY      841 AGCGCAAAATCAAGTCAAGTCAACCGGACAGAGGTGATGCTGTCCCTCATGGTAAACATTAACA 900
Db      841 AGCGCAAAATCAAGTCAAGTCAACCGGACAGAGGTGATGCTGTCCCTCATGGTAAACATTAACA 900

```

Qy	901	CTTTATCCCTTATGAACAAATGCTGAATTTGGAAAAACGAATTTGCTGTAATTATTTCCCCCT	960
Db	901	CTTTATCCCTTATGAACAAATGCTGAATTTGGAAAAACGAATTTGCTGTAATTATTTCCCCCT	960
Qy	961	TCGTTATCGTTTCAACCAATTTGGGTACAGATTTCAGACCAGAAACAACCAAGTCCCAATC	1020
Db	961	TCGTTATCGTTTCAACCAATTTGGGTACAGATTTCAGACCAGAAACAACCAAGTCCCAATC	1020
Qy	1021	GACTCCGGAACTTAGTCCAAAGTCGCAACCTGCACCAAAATTCCTCAACCAGTCCAAAGCAA	1080
Db	1021	GACTCCGGAACTTAGTCCAAAGTCGCAACCTGCACCAAAATTCCTCAACCAGTCCAAAGCAA	1080
Qy	1081	TCCAAATGATGAGAAATTTGGTCAAGAAGCTGTCGAAAAGTAGGCGATGGTTATGTCCT	1140
Db	1081	TCCAAATGATGAGAAATTTGGTCAAGAAGCTGTCGAAAAGTAGGCGATGGTTATGTCCT	1140
Qy	1141	TGAGGAAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAACCAACGACGC	1200
Db	1141	TGAGGAAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAACCAACGACGC	1200
Qy	1201	AGGCATTGATAGCAAACTGGCCCAAGCAGGAAGTTTATCTCATAGCTAGGAGCTAAGAA	1260
Db	1201	AGGCATTGATAGCAAACTGGCCCAAGCAGGAAGTTTATCTCATAGCTAGGAGCTAAGAA	1260
Qy	1261	AACAGCTCCCTCATCTAGTGATCGAGAAATTTTCAATAAGGCTTATGACTTACTAGCAAG	1320
Db	1261	AACAGCTCCCTCATCTAGTGATCGAGAAATTTTCAATAAGGCTTATGACTTACTAGCAAG	1320
Qy	1321	AATTCACCAAGATTTACTTGATATATAAGGTCGACAAAGTTGATTTGAGGCTTTGGATAA	1380
Db	1321	AATTCACCAAGATTTACTTGATATATAAGGTCGACAAAGTTGATTTGAGGCTTTGGATAA	1380
Qy	1381	CCTGTTTGGAAACGACTCAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTCTGGANGATTTCT	1440
Db	1381	CCTGTTTGGAAACGACTCAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTCTGGANGATTTCT	1440
Qy	1441	TGCGCTTCTTAGCTCCGATTCGTCATCCAGAAAGTTTGTAGAAAACCAAAATCGCGAAATTAC	1500
Db	1441	TGCGCTTCTTAGCTCCGATTCGTCATCCAGAAAGTTTGTAGAAAACCAAAATCGCGAAATTAC	1500
Qy	1501	CTACACTGATGATGCAATTCAGTAGCAGTTCGAGGAGGAGTACACAAACGAGACCG	1560
Db	1501	CTACACTGATGATGCAATTCAGTAGCAGTTCGAGGAGGAGTACACAAACGAGACCG	1560
Qy	1561	TTATATCTTTTGAATCTCGTGATATAACAGTGCATGAGGGGATGCTATGTAACCTCCACA	1620
Db	1561	TTATATCTTTTGAATCTCGTGATATAACAGTGCATGAGGGGATGCTATGTAACCTCCACA	1620
Qy	1621	TATGACCCATAGCCACTGGATTAAAAAGATGTTTGTCTGAAGCTGAGAGACGCGACG	1680
Db	1621	TATGACCCATAGCCACTGGATTAAAAAGATGTTTGTCTGAAGCTGAGAGACGCGACG	1680
Qy	1681	CCAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGG	1740
Db	1681	CCAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGG	1740
Qy	1741	AAATACTGAGGCAAAAAGGAGCAGACGACTATCTACAACCGCGTGAAGCAGCTTAAGAAGT	1800
Db	1741	AAATACTGAGGCAAAAAGGAGCAGACGACTATCTACAACCGCGTGAAGCAGCTTAAGAAGT	1800
Qy	1801	GCCACTTGATCGTATGACCATTAACATCTCAAAATTTGAGTGGTTGACGAAGGCGCT	1860
Db	1801	GCCACTTGATCGTATGACCATTAACATCTCAAAATTTGAGTGGTTGACGAAGGCGCT	1860
Qy	1861	AATCATACCTCATATGACCATTAACATCTCAAAATTTGAGTGGTTGACGAAGGCGCT	1920
Db	1861	AATCATACCTCATATGACCATTAACATCTCAAAATTTGAGTGGTTGACGAAGGCGCT	1920
Qy	1921	TTATGAGGCACCTTAAGGGGTACTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGT	1980
Db	1921	TTATGAGGCACCTTAAGGGGTACTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGT	1980
Qy	1981	CGAACTCCAAACGAACGTCGCAATTCAGATTAATGGTTTGTGTAACGCTAGCGACCATGT	2040

1981 CCAACATCCAAACGACGTCGCGCATTCAGATAATGTTTGGTAACGCTAGCGACCATGT 2040
2041 TCAAGAAACAAACAAATGGTCAAGCTGATCAACATCAACCGGAAACCAAGCGAGGAA 2100
2041 TCAAGAAACAAACAAATGGTCAAGCTGATCAACATCAACCGGAAACCAAGCGAGGAA 2100
2101 ACCTCAGACAGAAACCTTGAGGAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAA 2160
2101 ACCTCAGACAGAAACCTTGAGGAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAA 2160
2161 ACCAGAGTCTCAAAACCAACAGAGAAACCAAGAAATCAACAGAGAAATCAAGAGAAC 2220
2161 ACCAGAGTCTCAAAACCAACAGAGAAACCAAGAAATCAACAGAGAAATCAAGAGAAC 2220
2221 TCAGGTCGAGACTGAAAGAGTTCGAAGAAACCTGAGAGAGGCTGAAGATTTACTTGA 2280
2221 TCAGGTCGAGACTGAAAGAGTTCGAAGAAACCTGAGAGAGGCTGAAGATTTACTTGA 2280
2281 AATCCAGGAT 2290
2281 AATCCAGGAT 2290

RESULT 3

JS-09-468-656A-11
; Sequence 11, Application US/09468656A
; Patent No. 6582706
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-444
; CURRENT APPLICATION NUMBER: US/09/468,656A
; CURRENT FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
JS-09-468-656A-11

Query Match 97.0%; Score 2220.2; DB 4; Length 2531;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2253; Conservative 3; Mismatches 16; Indels 18; Gaps 1;
2Y 1 TTGTTCCATGAACTTGGTCGTCACCAAGCTGTGAGTTAAGAAAGCTTAATCGAGT 60
DB 57 TTGTTCCATGAGTCTGGACGTTTACCAAGCTGTGAGTTAAGAAAGCTTAATCGAGT 116
2Y 61 TTCTTATATAGATGGTATCAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 117 TGCTTATATAGATGGTATCAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 176
2Y 121 CAGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 177 CAGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 236
2Y 181 TTGACCTCTATGGAGACCATTAATCACTAATATGGAAGGAGGAGGAGGAGGAGGAGG 240
DB 237 TGTGACCTCTATGGAGACCATTAATCACTAATATGGAAGGAGGAGGAGGAGGAGGAGG 296
2Y 241 CATCAGTGAAGAGCTCTCATGAAAGATCCGAATTAATCACTAATATGGAAGGAGGAGG 300
DB 297 CATCAGTGAAGAGCTCTCATGAAAGATCCGAATTAATCACTAATATGGAAGGAGGAGG 356
2Y 301 CAATGAAATCAAGGGTGGTATGTCATTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTA 360

DB 357 CAATGAAATCAAGGGTGGTATGTCATTAAGGTAAACCGTAAATCTATGTTTACCTTAA 416
QY 361 GATGTCAGCTCATCGGATAATATTCGACAAAGAGAGATTAACGTCAGAGCAGGA 420
DB 417 GATGTCAGCTCATCGGATAATATTCGACAAAGAGAGATTAACGTCAGAGCAGGA 476
QY 421 AGCAGTCAATATCAATCAAGAGCAGATAATCTGTTGTCAGCAGAGCCCAAG 480
DB 477 AGCAGTCAATATCAATCAAGAGCAGATAATCTGTTGTCAGCAGAGCCCAAG 536
QY 481 ACCTTATACACGAGTATGATGATATCTTCAATGATGATATCAATGAGGACACGG 540
DB 537 ACCTTATACACGAGTATGATGATATCTTCAATGATGATATCAATGAGGACACGG 596
QY 541 TGATGCTTATATCTCTCAGCGGACCATTAACCATTAATCTTCAATGATGATATCA 600
DB 597 TGATGCTTATATCTCTCAGCGGACCATTAACCATTAATCTTCAATGATGATATCA 656
QY 601 AGCTAGCGAGTATGCTGTCAGAGAGCCCTATTGGAATGGAGAGGAGGATCTGCTCT 660
DB 657 AGCTAGCGAGTATGCTGTCAGAGAGCCCTATTGGAATGGAGAGGAGGATCTGCTCT 716
QY 661 TTCAAGTCTAGTATTAATGCAATCCAGCTCAACCAAGATTTGTCAGAGAACCAATCT 720
DB 717 TTCAAGTCTAGTATTAATGCAATCCAGCTCAACCAAGATTTGTCAGAGAACCAATCT 776
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACG 780
DB 777 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACG 836
QY 781 ATTGTATGCTAAACCTTATCAGAACGCTATGTAATCTGATGGCCTTTATTTTCGACC 840
DB 837 ATTGTATGCTAAACCTTATCAGAACGCTATGTAATCTGATGGCCTTTATTTTCGACC 896
QY 841 AGCGCAATCAAGAGTCGAAACCGCAGAGGTGAGTGTCTCTCATGTAACCATTAACA 900
DB 897 AGCGCAATCAAGAGTCGAAACCGCAGAGGTGAGTGTCTCTCATGTAACCATTAACA 956
QY 901 CTTTATCCCTTATGACAAATGTCGTAATGGGAAACGAAATTTGCTGCTTATTTCCCT 960
DB 957 CTTTATCCCTTATGACAAATGTCGTAATGGGAAACGAAATTTGCTGCTTATTTCCCT 1016
QY 961 TCGTATATGCTTCAACCAATTTGGTACAGATTCAGAACCAAGGAGGAGGAGGAGGAGG 1020
DB 1017 TCGTATATGCTTCAACCAATTTGGTACAGATTCAGAACCAAGGAGGAGGAGGAGG 1076
QY 1021 GACTCCGGAACCTAGTCCAGTCCGCAACCTGCAACCAATCTCAACCAAGCTCCAGCAA 1080
DB 1077 GACTCCGGAACCTAGTCCAGTCCGCAACCTGCAACCAATCTCAACCAAGCTCCAGCAA 1118
QY 1081 TCCAAATTTGATGAGAAATTTGGTCAAGAGCTGTTTCAAGAGTGGGAGTGGTATGCTTT 1140
DB 1119 TCCAAATTTGATGAGAAATTTGGTCAAGAGCTGTTTCAAGAGTGGGAGTGGTATGCTTT 1178
QY 1141 TGAGAGAGTGGAGTGTCTCGTTATATCCAGCAAGGATCTTTCCAGCAAGAACGAGCAGC 1200
DB 1179 TGAGAGAGTGGAGTGTCTCGTTATATCCAGCAAGGATCTTTCCAGCAAGAACGAGCAGC 1238
QY 1201 AGGCAITTGATGCAAACTGGCCCAAGAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260
DB 1239 AGGCAITTGATGCAAACTGGCCCAAGAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1298
QY 1261 AACTGACCTCCCATCTAGTATGATGAGAAATTTTACAATAGGCTTATGATCTACTAGCAAG 1320
DB 1299 AACTGACCTCCCATCTAGTATGATGAGAAATTTTACAATAGGCTTATGATCTACTAGCAAG 1358
QY 1321 AATTCAACCAAGATTTTACTGATATTAAGGTCGACAGCTTCTGATTTTGGGCTTTGGATA 1380
DB 1359 AATTCAACCAAGATTTTACTGATATTAAGGTCGACAGCTTCTGATTTTGGGCTTTGGATA 1418
QY 1381 CCTGTTGGAACGACTCAAGGATGTCNCAAGGTGATAAAGTCAAGTGTAGTGGAAGATTTCT 1440
DB 1419 CCTGTTGGAACGACTCAAGGATGTCNCAAGGTGATAAAGTCAAGTGTAGTGGAAGATTTCT 1478

897 CCAGCGCAATCACAAAGTCGAAACCGCCAGAGGTGTAGCTGTCCCTCATGTAAACCAATTA 956
898 CCACATTTATCCCTTATGAAACAAATGCTGAATTCGAAACAAAGATTCGCTGATTTATTC 957
957 CCACATTTATCCCTTATGAAACAAATGCTGAATTCGAAACAAAGATTCGCTGATTTATTC 1016
958 CTTTCGTTATCGTTTAAACCAATTCGTTGTTACAGATTCAGATTCAGATTCAGATTCAG 1017
1017 CTTTCGTTATCGTTTAAACCAATTCGTTGTTACAGATTCAGATTCAGATTCAGATTCAG 1076
1018 ATGAGCTCCGGAACCTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCAG 1077
1077 ATGAGCTCCGGAACCTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCAG 1136
1078 CAATCCAAATGATGAGAAATTCGTTCAAGAGAGCTGTTGAAAGATGAGCGATGTTATGT 1137
1137 CAATCCAAATGATGAGAAATTCGTTCAAGAGAGCTGTTGAAAGATGAGCGATGTTATGT 1196
1138 CTTTGAGGAAATGAGATTCCTGTTATATCCAGCAAGATTCCTTCAGCAGAAACAGC 1197
1197 CTTTGAGGAAATGAGATTCCTGTTATATCCAGCAAGATTCCTTCAGCAGAAACAGC 1256
1198 AGCAGGATTCAGTAAACCTGCGCAAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTAA 1257
1257 AGCAGGATTCAGTAAACCTGCGCAAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTAA 1316
1258 GAAATCTGACCTCCCATCTAGTATGATGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1317
1317 GAAATCTGACCTCCCATCTAGTATGATGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1376
1318 AAGAAATCCACGAATTTACTTGAATAAAGTTCGACAAAGTTCGATTTGAGGCTTTGGA 1377
1377 AAGAAATCCACGAATTTACTTGAATAAAGTTCGACAAAGTTCGATTTGAGGCTTTGGA 1436
1378 TAACTGTTGGAACGACTCAAGATGTCNCAAGTATAAAGTCAAGTATGAGGATGATAT 1437
1437 TAACTGTTGGAACGACTCAAGATGTCNCAAGTATAAAGTCAAGTATGAGGATGATAT 1496
1438 TCTGCTCTTCTAGTCCGATTCGATTCATCAGAACGTTTGAAGAACCAATTCGCAAT 1497
1497 TCTGCTCTTCTAGTCCGATTCGATTCATCAGAACGTTTGAAGAACCAATTCGCAAT 1556
1498 TACCTACTGATGATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1557
1557 TACCTACTGATGATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1616
1558 CGTTATATCTTGTATCTGCTGATATACCAAGTATGAGGAGGATGCTTATGATTC 1617
1617 CGTTATATCTTGTATCTGCTGATATACCAAGTATGAGGAGGATGCTTATGATTC 1676
1618 ACATATGACCCATAGCCACTGGAATTAAGAAAGATGTTGCTGAGAGGAGGAGGCGC 1677
1677 ACATATGACCCATAGCCACTGGAATTAAGAAAGATGTTGCTGAGAGGAGGAGGCGC 1736
1678 AGCCAGGCTTATGCTTAAAGAAAGGTTTGACCCCTCTTCACAGACCATCAGATTC 1737
1737 AGCCAGGCTTATGCTTAAAGAAAGGTTTGACCCCTCTTCACAGACCATCAGATTC 1796
1738 AGGAATATCTAGGCAAAAGGAGCAGAGCTATCTACACCGGCTGAAAGCAGCTAAGAA 1797
1797 AGGAATATCTAGGCAAAAGGAGCAGAGCTATCTACACCGGCTGAAAGCAGCTAAGAA 1856
1798 GGTGCACTGATGCTGATGCTTACATCTTCAATATCTGTAAGATCAAAACCGTAG 1857
1857 GGTGCACTGATGCTGATGCTTACATCTTCAATATCTGTAAGATCAAAACCGTAG 1916
1858 TTTAATCATACCTCATATGACCATTAACATCAATCAATTTAGTGTGTTGACCAAG 1917
1917 TTTAATCATACCTCATATGACCATTAACATCAATCAATTTAGTGTGTTGACCAAG 1976
1918 CTTTATGAGGCAACCTAAGGAGTATCTCTGAGGATCTTTGCGGACTGTCAGTACTA 1977

Db 1977 CTTTATGAGGCAACCTAAGGAGTATCTCTGAGGATCTTTTGGCGACTGTCAGTACTA 2036
QY 1978 TGTGAAATCAATCAAAACGAAAGTCCGATTCAGATTAATGTTTGGTAAACGCTAGCGACCA 2037
Db 2037 TGTGAAATCAATCAAAACGAAAGTCCGATTCAGATTAATGTTTGGTAAACGCTAGCGACCA 2096
QY 2038 TGTTCAAAGAAACAAAAATGTTCAAGCTGATACCAATCAAAACGAAAGGAGGAGGAGG 2097
Db 2097 TGTTCGTAATAATAAGGTAGACCAAGACAGTAAACCTGATGAAGATAAGGAACATGATGA 2156
QY 2098 GAAACCTCAGACAGAAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2133
Db 2157 AGTAAGTACGAGCAACTCACCTGATGATGAAAGAGAGATCAAGCTGTTTAAATCC 2216
QY 2134 TCGAGAGAGAGAAACCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2193
Db 2217 TTAGCAGATTAATCTTTTAAACCAAGCTGATACGGAAGAGACAGAGGAGGAGGAGGAGG 2276
QY 2194 AGAATCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2253
Db 2277 AGATACCAACAGATGAGGCTGAAATTCCTCAAGTAGAGATTCGTTATTAACGCTAAGAT 2336
QY 2254 GAGAGAGGCTGAAGATTTACTTTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2290
Db 2337 AGCAGATCGGAGGCTTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2373

RESULT 5

US-08-961-527-94
; Sequence 94, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-94

Query Match 87.3%; Score 1999.2; DB 4; Length 8195;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 2135; Conservative 1; Mismatches 154; Indels 27; Gaps 2;

QY 1 TTGTTCTTATGAACCTTGCTGTCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAAATCGAGT 60
Db 5658 TTGTTCTTATGAACCTTGCTGTCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAAATCGAGT 5717
QY 61 TTCTTATATAGATGGTATCAGGCTGGTCAAAAGCGCAGAAAACCTTGACACCAAGATGAAGT 120
Db 5718 TKCTTATATAGATGGTATCAGGCTGGTCAAAAGCGCAGAAAACCTTGACACCAAGATGAAGT 5777
QY 121 CAGTAAGAGGGAGGGATCAAGCGCGCAACAAATNGTNTATCAAGATTACGGATCAAGGTTA 180
Db 5778 CAGTAAGAGGGAGGGATCAAGCGCGCAACAAATCGTCATCAAGATTACGGATCAAGGTTA 5837
QY 181 TGTGACCTCTCATGAGACCAATATCATTAATTAATGSCAAGGTTCTTATGATGCCAAT 240
Db 5838 TGTGACCTCTCATGAGACCAATATCATTAATTAATGSCAAGGTTCTTATGATGCCAAT 5897
QY 241 CATCAGTCAAGAGCTCCATCAAGAGATCCGAATTTATCAGTTGGAAGATTACAGATTTGT 300
Db 5898 CATCAGTCAAGAGCTCCATCAAGAGATCCGAATTTATCAGTTGGAAGATTACAGATTTGT 5957
QY 301 CAATCAAAATCAAGGTTGTTATGTCAATTAAGGTAACCGTAAATATCTATGTTACCTTAA 360
Db 5958 CAATCAAAATCAAGGTTGTTATGTCAATTAAGGTAACCGTAAATATCTATGTTACCTTAA 6017
QY 361 GGATGACGCTCATGCGGATATATTTGGACAAAAGAGAGATTAAACGTCAGAACGAGA 420
Db 6018 GGATGACGCTCATGCGGATATATTTGGACAAAAGAGAGATTAAACGTCAGAACGAGA 6077
QY 421 ACGCAGTCAATATCAACTCAAGAGC---AGATAATGCTGTTGCTGACGACAGGCCCA 477
Db 6078 ACAAGTCAATATCAACTCAAGAGC---AGATAATGCTGTTGCTGACGACAGGCCCA 6137
QY 478 AGGACGTTATCAACGGATGATGGTATATCTTCAATGCTGATCTGATATCATTTAGGACAC 537
Db 6138 AGGACGTTATCAACGGATGATGGTATATCTTCAATGCTGATCTGATATCATTTAGGACAC 6197
QY 538 GGGTGATGCTTATATCGTTCTCAAGGACCAATACCAATTAATTCCTTAAAGATGAGTT 597
Db 6198 GGGTGATGCTTATATCGTTCTCAAGGACCAATACCAATTAATTCCTTAAAGATGAGTT 6257
QY 598 ATCAGTAGCAGGTTAGTCTGTCAGAGAGCCTATTGGAAATGGGAAAGCAGGATCTGTC 657
Db 6258 ATCAGTAGCAGGTTAGTCTGTCAGAGAGCCTATTGGAAATGGGAAAGCAGGATCTGTC 6317
QY 658 TTCTTCAAGTTCTAGTTATATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCA 717
Db 6318 TTCTTCAAGTTCTAGTTATATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCA 6377
QY 718 TCTGACTGTCACTCAACTTATCATCAAAAATCAAGGGGAAACATTTCAAGCTTTTACG 777
Db 6378 TCTGACTGTCACTCAACTTATCATCAAAAATCAAGGGGAAACATTTCAAGCTTTTACG 7437
QY 778 TGAATTTGATGCTAAACCTTATCAGAGGCGCATGTGGAATCTGATGGCCTTATTTTCA 837
Db 6438 TGAATTTGATGCTAAACCTTATCAGAGGCGCATGTGGAATCTGATGGCCTTATTTTCA 6497
QY 838 CCCAGCGCAATTCAGAGTCAAGACCGCAGAGGCTGTAGTGTCCCTCATGTTAAACATTA 897
Db 6498 CCCAGCGCAATTCAGAGTCAAGACCGCAGAGGCTGTAGTGTCCCTCATGTTAAACATTA 6557
QY 898 CCACCTTTATCCCTTATGAACAAATGCTGAATTTGGAATAACGAATGCTGCTATTATTC 957
Db 6558 CCACCTTTATCCCTTATGAACAAATGCTGAATTTGGAATAACGAATGCTGCTATTATTC 6617
QY 958 CTTTGGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACAGAACCAACCAAGTCCACA 1017
Db 6618 CTTTGGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACAGAACCAACCAAGTCCACA 6677
QY 1018 ATCGACTCGGNAACCTAGTCCAGTCCGCACTGCAACCAATTCCTCAACAGCTCCCAAG 1077
Db 6678 ATCGACTCGGNAACCTAGTCCAGTCCGCACTGCAACCAATTCCTCAACAGCTCCCAAG 6737
QY 1078 CAATCCAAATTTGATGAGAAATTTGGTCAAGAACGCTGTTCGAAAAGTAGGCGATTATGT 1137

Db 6738 CAATCCAAATTTGATGAGAAATTTGGTCAAGAAAGCTGTTCCAAAAGTAGGCGATGTTATGT 6797
QY 1138 CTTTGGAGGAGAAATGGAGTTTCTCGTTATATATCCAGCCAAAGGATCTTTTCAGCAGAAAACAGC 1197
Db 5798 CTTTGGAGGAGAAATGGAGTTTCTCGTTATATATCCAGCCAAAGGATCTTTTCAGCAGAAAACAGC 6857
QY 1198 AGCAGGCATTTGATAGCAAACTGGCCAAAGCAGGAGAAAGTTTATCTCATAGCTTAGGAGCTAA 1257
Db 6858 AGCAGGCATTTGATAGCAAACTGGCCAAAGCAGGAGAAAGTTTATCTCATAGCTTAGGAGCTAA 6917
QY 1258 GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1317
Db 6918 GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 6977
QY 1318 AAGAAATTCACCAAGATTCTTCTGATATAAAGGTCGACAAAGTTGATTTTTCAGGCTTTTGA 1377
Db 6978 AAGAAATTCACCAAGATTCTTCTGATATAAAGGTCGACAAAGTTGATTTTTCAGGCTTTTGA 7037
QY 1378 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATAT 1437
Db 7038 TAACCTGTTGGAAACGACTCAAGGATGTCYCAAGTGTATAAAGTCAAGTTAGTGGANGATAT 7097
QY 1438 TCTTGGCTTCTTAGCTCCGATTCGATCCAGAACGTTTATAGGAAAAACCAATGCGCAAT 1497
Db 7098 TCTTGGCTTCTTAGCTCCGATTCGATCCAGAACGTTTATAGGAAAAACCAATGCGCAAT 7157
QY 1498 TACCTACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGA 1557
Db 7158 TACCTACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGA 7217
QY 1558 CGGTTATATCTTTGATCTCGTGATATACAGTGTATGAGGGGATCGCTATGTAATCC 1617
Db 7218 CGGTTATATCTTTGATCTCGTGATATACAGTGTATGAGGGGATCGCTATGTAATCC 7277
QY 1618 ACATATGACCCCATAGCCACTGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGC 1677
Db 7278 ACATATGACCCCATAGCCACTGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGC 7337
QY 1678 AGCCGAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCAATCAGGATTC 1737
Db 7338 AGCCGAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCAATCAGGATTC 7397
QY 1738 AGGAATACTGAGGCAAAAGGACAGACAGTATCTACAAACCGGCTGAAGCAGCTAAGAA 1797
Db 7398 AGGAATACTGAGGCAAAAGGACAGACAGTATCTACAAACCGGCTGAAGCAGCTAAGAA 7457
QY 1798 GGTGCCACTTTGATCGTATGCTTACCAATCTTCAATATCTGTAGAAGTCAAAAAACGGTAG 1857
Db 7458 GGTGCCACTTTGATCGTATGCTTACCAATCTTCAATATCTGTAGAAGTCAAAAAACGGTAG 7517
QY 1858 TTTAATCATACCTCATTTATGACCAATTCACCAATACCAATTTAGTGGTTTGAAGG 1917
Db 7518 TTTAATCATACCTCATTTATGACCAATTCACCAATACCAATTTAGTGGTTTGAAGG 7577
QY 1918 CTTTTATGAGGCACTTAAAGGGTATCTCTTGAAGTCTTTTGGCAGCTCTCAAGTACTA 1977
Db 7578 CTTTTATGAGGCACTTAAAGGGTATCTCTTGAAGTCTTTTGGCAGCTCTCAAGTACTA 7637
QY 1978 TGTGCAACATCCAAAACGACCTCCGATTCAGATTAATGGTTTTTGGTAAACGCTACCGACCA 2037
Db 7638 TGTGCAACATCCAAAACGACCTCCGATTCAGATTAATGGTTTTTGGTAAACGCTACCGACCA 7697
QY 2038 TGTTCAGAGAAACAAAATGGTCAAGCTGATACCAATCAAAACGGGAAAAACCAAGCAGGA 2097
Db 7698 TGTTCAGAGAAACAAAATGGTCAAGCTGATACCAATCAAAACGGGAAAAACCAAGCAGGA 7757
QY 2098 GAAACCTCAGACAGAAAACCTCGAGGAAG-----AACCCC 2133
Db 7758 AGTAAGTGAAGCAACTCACCTGAAATCTGATGAAAAAGAGAAATCACGCTGTTTAAATCC 7817
QY 2134 TCAGAGAGAGAAACCGCAAGCGAGAGAAACCAAGATGTCCTCAAAACCAACAGAGGAACAGGA 2193

7818 TTACGACAGATTAATCTTTTATATAAACCGACACTGATACGAGACAGACAGAGAGAACTGA 7877

2194 AGAATCACGAGGGAATCAGAGAAACCTCAGCTCGAGACTGAAAGGTTTGAGAGAAAACT 2253

7878 AGATACCACAGATGAGGCGTGAATTCCTCAGTAGAGAAATCTGTTATTAAACGCTAAGAT 7937

2254 GAGAGGGCTCAGGATTTACTTGGAAAAATCCAGGAT 2290

7938 AGCAGATCGGAGGCGCTTGTGAGAAAAAGTAAACAGAT 7974

RESULT 6

JS-08-961-527-243
 / Sequence 243 Application US/08961527
 / Patent No. 6420135
 / GENERAL INFORMATION:
 / APPLICANT: Charles Kunath
 / TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 / NUMBER OF SEQUENCES: 391
 / CORRESPONDENCE ADDRESS:

DATE: 20830
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

SOFTWARE: ASCII TEX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

Query Match	60.1%	Score 1376.6	DB 4	Length 2359
Best Local Similarity	97.8%	Pred. No. 0		
Matches 1410	Conservative 0	Mismatches 14	Indels 18	Gaps 1
1	TTGTTCCTATGAACTTGGTCGTGCATCACCAGCTGGTCAGGTTAAGAAAGAGCTCAATCGAGT	60		
936	TTGTTCCTATGAGCTTGGAGCTTACCAAGCTGGTCAGGATPAGAAAGAGTCTCAATCGAGT	995		
61	TTCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCAGAAAACTTCACACCATCAAGT	120		
996	TGCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCAGAAAACTTCACACCATCAAGT	1055		
121	CAGTAACAGAGGAGGGGATCAACGCCGAACAAATNGTNATCAAGATTACGGATCAAGGTTA	180		
1056	CAGTAACAGAGGAGGGGATCAACGCCGAACAAATNGTNATCAAGATTACGGATCAAGGTTA	1115		
181	TGTGACCTCTCATGGAGACCATTAATCAATTAATATGCGAAGGTTCTTTATGATGCCAT	240		
1116	TGTGACCTCTCATGGAGACCATTAATCAATTAATATGCGAAGGTTCTTTATGATGCCAT	1175		

QY	241	CATCAGTGAAGAGCTCCTCATGAAAGAGATCGGAATTTATCAGTTGAAGATTTCAGACATTGT	300
DB	1176	CATCAGTGAAGAGCTCCTCATGAAAGATCGGAATTTATCAGTTGAAGATTTCAGACATTGT	1235
QY	301	CAATGAAATCAAGGGGTGTTATGTCATTAAAGTAAACGGTAATAACTATCTGNTACCTTAA	360
DB	1236	CAATGAAATCAAGGGGTGTTATGTCATTAAAGTAAACGGTAATAACTATCTGNTACCTTAA	1295
QY	361	GGATGCAGCTCATGCGGATAATATTCCGACAAAGAGAGATTAAACGTCAGAAAGCAGGA	420
DB	1296	GGATGCAGCTCATGCGGATAATATTCCGACAAAGAGAGATTAAACGTCAGAAAGCAGGA	1355
QY	421	ACGCAGTCAATAATCAATACTCAAGAGCAGATATGCTGTTGCTGACCCAGAGGCCAAGG	480
DB	1356	ACGCAGTCAATAATCAATACTCAAGAGCAGATATGCTGTTGCTGACCCAGAGGCCAAGG	1415
QY	481	ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTACAGACACGGG	540
DB	1416	ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTACAGACACGGG	1475
QY	541	TGATGCTTATATCGTTCTCTCAGCGGACCAATTACCAATTAATTCCTAAGAATGAGTTATC	600
DB	1476	TGATGCTTATATCGTTCTCTCAGCGGACCAATTACCAATTAATTCCTAAGAATGAGTTATC	1535
QY	601	AGCTAGCGATTAGCTGCTCAGAAAGCTTATGGAATGGGAAGCAGGATCTCGTCTTC	660
DB	1536	AGCTAGCGATTAGCTGCTCAGAAAGCTTATGGAATGGGAAGCAGGATCTCGTCTTC	1595
QY	661	TTCAAGTCTAGTATTAATGCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	720
DB	1596	TTCAAGTCTAGTATTAATGCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	1655
QY	721	GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	780
DB	1656	GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	1715
QY	781	ATTGTATGTAAACCCCTTATCAGAACGCCATGTGGAATCTGATGGGCTTATTTTCGACCC	840
DB	1716	ATTGTATGTAAACCCCTTATCAGAACGCCATGTGGAATCTGATGGGCTTATTTTCGACCC	1775
QY	841	AGCGCAATCACAAGTTCGAACCGCCAGAGGTGTAGCTGTCCCTCATGTTAACCATTCAC	900
DB	1776	AGCGCAATCACAAGTTCGAACCGCCAGAGGTGTAGCTGTCCCTCATGTTAACCATTCAC	1835
QY	901	CTTTATCCCTTATGAACAAATGTCTGAAATGGAAAAACGAATTCGCTGATTTATTCCTCT	960
DB	1836	CTTTATCCCTTATGAACAAATGTCTGAAATGGAAAAACGAATTCGCTGATTTATTCCTCT	1895
QY	961	TCGTTATCGTTCAAAACCATTTGGGTACAGATTCAAGACCAAGTCCCAATC	1020
DB	1896	TCGTTATCGTTCAAAACCATTTGGGTACAGATTCAAGACCAAGTCCCAATC	1955
QY	1021	GACTCGGAACCTTAGTCCCAAGTCCGCAACCTGCACCAATCTCTCAACCGAGCTCCAGCA	1080
DB	1956	GACTCGGAACCTTAGTCCCAAGTCCGCAACCTGCACCAATCTCTCAACCGAGCTCCAGCA	1997
QY	1081	TCCAAATGATGAAATTCGTTCAAAGAACTGTTTCGAAAAGTAGGCGATGGTTATCTCTT	1140
DB	1998	TCCAAATGATGAAATTCGTTCAAAGAACTGTTTCGAAAAGTAGGCGATGGTTATCTCT	2057
QY	1141	TGAGGAAATGGAGTTTCTCGTTATATCCAGCCAGGATCTTTTCAGCAGAAACAGCAGC	1200
DB	2058	TGAGGAAATGGAGTTTCTCGTTATATCCAGCCAGGATCTTTTCAGCAGAAACAGCAGC	2117
QY	1201	AGGCATTGATAGCAAACTCGGCCAAGCAGAAAGTTTATCTCATAGCTTAGGAGCTTAAGAA	1260
DB	2118	AGGCATTGATAGCAAACTCGGCCAAGCAGAAAGTTTATCTCATAGCTTAGGAGCTTAAGAA	2177
QY	1261	AACGTACCTCCCATCTAGTGTGCGAAATTTTACAAATAGGCTTATGCTACTACTGCAAG	1320
DB	2178	AACGTACCTCCCATCTAGTGTGCGAAATTTTACAAATAGGCTTATGCTACTACTGCAAG	2237
QY	1321	AAPTCACCAGAATTTACTTTGATAATAAGAGTCGAACAAAGTTGATTTTGAAGCTTTGGATAA	1380

Db 2238 AATTCAACAGATTACTTGTATTAATTAAGGTGCAACAGTTGATTTTGGGCTTTGGATAA 2297
Qy 1381 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGCATAAAGTCAAGTTAGTGGANGATATTTCT 1440
Db 2298 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGCATAAAGTCAAGTTAGTGGANGATATTTCT 2357
Qy 1441 TG 1442
Db 2358 TG 2359

RESULT 7
US-09-468-656A-9
; Sequence 9, Application US/09468656A
; Patent No. 6582706
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-444
; CURRENT APPLICATION NUMBER: US/09/468,656A
; CURRENT FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2451)
; OTHER INFORMATION: n = a, c, t or g

US-09-468-656A-9

Query Match 43.3%; Score 990.6; DB 4; Length 2451;
Best Local Similarity 67.8%; Pred. No. 1e-293;
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

Qy 1 TTGTTCTATGAATCTGCTGCTCACCAGCTGGTCAAGTTAAGAAAGAGTCTTAATCGAGT 60
Db 57 TTGTTCTATGAGTTGGAGCTGTATCAAGCTAGAACGGTAAAGGAAA--TAATCGGT 113
Qy 61 TTCTTATATAGATGGTATCAGGCTGGTCAAAAGCGAAGAACTTGACACAGATGAAGT 120
Db 114 TTCTTATATAGATGAAACAAAGCGACGCAAAACAAACCGAGAAATTTGACTCTCTGATGAGT 173
Qy 121 CAGTAAGAGGGAGGGATCAACGCGGAACAAATGNTATCAAGATTACGGATCAAGGTTA 180
Db 174 TAGCAAGCGTGAAGGAATCATGCTGAGCAATCGTCAACAGATAACAGACCAAGGCTA 233
Qy 181 TGTGACCTCTCATGAGACCATTAATCAATTAATGATGCAAGGTTCCCTTATGATGCCAT 240
Db 234 TGTCACTTCAATGGCGACCACTATCAATTAATCAATGCTAAGGTTCTTATGACGCTAT 293
Qy 241 CATCAGTGAAGAGCTCTCTCATGAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTG 300
Db 294 CATCAGTGAAGAAATTAATCTCATGAAGATCCGAATTAATGAAGCTAAGAAATGAGGATTTGT 353
Qy 301 CAATGAATCAAGGGTGGTATGATCAATTAAGTAAACGGTAAATCAATGNTACTTAA 360
Db 354 TAATGAGTCNAGGGTGGATGATGTTATCAAGGTAGATGGAATACTATGTTTACCTTAA 413
Qy 361 GGATGACGCTCATGCGGATTAATTTGACACAAAAGAGATTAACACGTCGAGACGAGA 420
Db 414 GGATGCTGCCACGGGATACGTCGGTACAAAGAGGAATCAATCGACAAAACAGA 473
Qy 421 ACGCAGTCATAATCAT-----AACTCAAGAGCGAGATAATGCTGTTGTCAGGCCAG 471

Db 474 GCATAGTCAACATCGTGAAGGTGGAACCTCCAGAGAACGATGGTGTCTGTTGCCTTGGACAG 533
Qy 472 AGCCCAAGGACGTTATATACAAACGGATGATGGGTATATCTTCAATGATCATCTGATATCATTTGA 531
Db 534 TTGCGAAGGACGCTATATACAGATGATGGTTATATCTTTAATGCTTCTGATATCATAGA 593
Qy 532 GGACACGGGTGATGCTTATATATCTTCTCAGCGGACCAATTAACATTAATTCCTTAAGAA 591
Db 594 GGATACGCTGATGCTTATATCTTCTCAGCGGATCAATTAACATTAATTCCTTAAGAA 653
Qy 592 TGAGTTATCAGCTAGCGGATTAAGTCTGTCAGAGCCCT----- 629
Db 654 TGAGTTATCAGCTAGCGGATTAAGTCTGTCAGAGCCCTTCTATCTGCTGAGGAAATCT 713
Qy 630 -----ATTGGAATGGGAAGCAGGATCTCGTCCCTTCTTCAAGTCTTAGTTATAA 678
Db 714 GTCAAAATTCAGAAACCTATCCCGGACAAATAGCGATAACACATTCAGAGAACAACTGGGT 773
Qy 679 TGAATAATCCAGCTCAACCAAGATTGTGAGAGAACCAATCTGATCTGCTCACTCCAACTTA 738
Db 774 ACCTTCTGTAAGCAATCCAGGAACCTACAAATATAACACAAAGCAACCAACCAACTAA 833
Qy 739 TCATCA---AAATCAAGGGGAAACAAATTTCAAGCCTTTTACGTAATGTATGTAAC 795
Db 834 CAGTCAAGCAAGTCAAGATTAATGACATTTGATCTCTTGAAACAGCTCTACAAACTGCC 893
Qy 796 CTATATCAGAACGCGCATGTGGAATCTGATGCGCTTATTTTCGACCGAGCAAAATCAACAAG 855
Db 894 TTTGAGTCAAGCATGTAGATCTGATGCGCTTGTCTTGATCCAGACAAATATCAACAG 953
Qy 856 TCGAACCGCAGAGGTGATGCTCCCTCATGTGTAACCATTAACCATTTATCCCTTATGA 915
Db 954 TCGAACAGCTAGAGGTGTCAGTGCACACAGGAGATCAATCCACATTCATCCCTTACTC 1013
Qy 915 ACAATGCTGTAATTTGAAAGCAAGATTTGCTGATATTTTCCCTTCTGTTATCGTTCAAA 975
Db 1014 TCAAAATGCTGAAATTTGAAAGCAAGATTCGTCGTATTTATTTCCCTTCTGTTATCGTTCAAA 1073
Qy 976 CATTGGGTACAGATTCAAGACCAAGCAACCAAGTCCACAATCGACTCGGAAACCTAG 1035
Db 1074 CATTGGGTACAGATTCAAGGCGCAGAACCAACAGTCCACAACCGACTCCGGAACCTAG 1133
Qy 1036 TCCAAAGTCGCAACCTGCAACCAATCTTCAACAGCTCCAAAGCAATCCAATTTGATGAA 1095
Db 1134 TCCAGGCCCGCAACCTGCAACCAATCTTAAATPAGACTCAA-----ATTCTTC 1181
Qy 1096 ATTGTCAAAGAGCTGTTGCAAAAGTAGCGATGTTATGCTTTGAGAGAGATGAGT 1155
Db 1182 TTTGGTTAGTCAAGTGGTACGAAAGTTGGGGAAGGATATGATTCGAGAAAGGGCAT 1241
Qy 1156 TTTCTGTTATATCCAGCCAGCATCTTTTCAGCAGAAACAGCAGCAGGCAATTCATGACAA 1215
Db 1242 CTCTCGTTATGCTTTGCGAAGATTTACATCTGAACTGTTAAATCTTGAAGCA 1301
Qy 1216 ACTGCGCAAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTTAAGAAACCTGACCTCCCATC 1275
Db 1302 GTTATCAAAACAGAGAGTGTTCACACTTTAATGCTTAAAGAAAGAAATGTTGCTCC 1361
Qy 1276 TAGTGAATCAGAAATTTTCAATTAAGCTTATGATCTTAGCAGAAATTCACCAAGATTT 1335
Db 1362 TCGTGACCAAGAAATTTTATGATAAAGCATATAATCTGTTAACTGAGGCTCATAAAGCCTT 1421
Qy 1336 ACTTCATATAAAGTCCGAAAGTTGATTTTTCAGGCTTTGGATAACCTGTTGAAACGACT 1395
Db 1422 GTTTGNAATAAGGGTCGTAATCTTGATTTCCAGGCTTTAGACAAATATTAGAACGCTT 1481
Qy 1396 CAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGANGATATTTTTCCTTCTTAGCTCC 1455
Db 1482 GAATGATGAATCGACTAATAAGAAATAATGATGATGATGATTTATTTGCGCATTCCTAGCAC 1541
Qy 1456 GATTCGTATCAGAAACGTTTATAGGAAACCAAAATGCGAAATTCCTACCTGATGATGA 1515
Db 1542 AATTACCATCCAGAGCGACTTTGGCAACCAAAATCTCAAAATTTGAGTATATCTAGTACTAGAGCA 1601

1516 GATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAAAGACGGTTATATCTTTGTATCC 1575
1602 AGTTGCTGTTGTTCAATAGCTGATAGTATACACAGCTCAGATGGTTATATTTGATGA 1661
1576 TCGTGATATACAGATGATGAGGGGATGCTATGTAACTCCACATATGACCCATAGCCA 1635
1662 ACATGATATATATCAGTATGATGAGGAGATGATATGTAACGCTCATATGGGCCATAGTCA 1721
1636 CTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCCAGCGCTTATGCTAA 1695
1722 CTGATTGGAAGATAGACCTTTCTGATAGGAAAGATGTCAGCTCAAGCCTATATCTAA 1781
1696 AGAGAAAGGTTTGACCCCTCTCTCCAGACAGCATCAGGATTACAGAAATATCTGAGGCAA 1755
1782 AGAAAAGGTTTCTACTCTCCATCTCCAGAGCGAGATGTTAAAGCAATCAACTGGAGA 1841
1756 AGAGCAGAAAGCTATCTACACCGCGTGAAGCAGCTAAGAAAGTGCCACTTGTATCGPAT 1815
1842 TAGTCAGCAGCTATTTACAACTGCTGTGAAGGGGAAAAACGAATTCACCTCGTTCCACT 1901
1816 GCCTTCAATCTTCAATATCTAGTAAAGTCAAAAGCGGTAGTTTAACTCATCTCATTA 1875
1902 TCCATATATGTTGAGCATACAGTTGAGGTAAAGCGGTAAAGCGGTAAATTTGATTTCTCTCAA 1961
1876 TGACCATTAACATAACATCAAAATTTGAGTGGTTTGAAGGCGCTTTATGAGGCACTTAA 1935
1962 GGATCATTACCATATATTAATTTGCTTGGTTGATGATCACATACAAAGCTCCAAA 2021
1936 GGGGTATCTTTGAGGATCTTTTGGCGACTGTCAGTACTATCTCGAATCTCAAAACGA 1995
2022 TGCGTATACCTTGAAGATTTGTTGCGAGCTTAAGTACTACGTAGAACACCTTGACGA 2081
1996 ACCTCCGATTCAGATATGTTTGTGAAGCTTACGAGCATGTTTCAAGAAACAAAA 2055
2082 ACCTCCACATCTTAATGATGATGGGCAATGCGCAGTGAGCATGTGTTAGGCAAGAA 2141
2056 TGCTCAAGCTGATACCAATCAAGCGAAACCAAGCGAGGAGAACTCTCAGACAGAAA 2115
2142 CCACAGTGAAGTCCAAATTAAGAACTTCAAGCGGATGAAGAGCCAGTAGAGGAAACACC 2201
2116 ACCTGAGGAAGAAACCCCTCCAGAAAGAGA 2144
2202 TGCTGAGCCAGAGTCCCTCAAGTAGAGA 2230

RESULT 8
US-09-765-272a-65
Sequence 55, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 2389 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
US-09-765-272a-65

Query Match 43.1%; Score 987.6; DB 3; Length 2389;
Best Local Similarity 67.7%; Pred. No. 8.3e-293;
Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

QY 4 TTCTTATGAACCTTGGTCCGTCACCAAGCTGTCAGAGTTTAAAGAGAGTCTAATCGAGTTTC 63
Db 1 TTCTTACGAGTTGGGACTGTATCAAGCTAGAACGGTTTAAAGGAAA---TAACTGTTTC 57
QY 64 TTATATAGATGCTGATCAGGCTCGTCAAAAGGCGAGAACTTCACACACAGATGAAGTCAG 123
Db 58 CTATATAGATGGAACCAAGCGACGCAAAACCGGAGATTTGACTCTCTGATGAGGTAG 117
QY 124 TAAGAGGAGGGGATCAAGCCGCAAAATNGTATCAAGATTACCGATCAAGGTTATGT 183
Db 118 CAAGCGTGAAGGAATCAATGCTGAGCAATCGTCATCAAGATAACAGACCAAGGCTATGT 177
QY 184 GACCTCTCATGAGACCATTAATCAATCAATGATGCAAGGTTCCCTTATGATGCCATCAT 243
Db 178 CACTTCATGATGCGACCATCTATCAATTAATGATGAGGTTCCCTTATGACGCTATCAT 237
QY 244 CAGTCAAGAGCTCTCATGAAAGATCCGAATATATCAGTTGAAGGATTACACATTTGTCAA 303
Db 238 CAGTGAAGATTTACTCTGAAAGATCCAACTATATAGCTAAAGATGAGATATTTGTAA 297
QY 304 TGAATCAAGGTTGTTATGTCTATTAAGGTAAACGGTAAATATCTATGTTNACCTTAAGGA 363
Db 298 TGAGGTCAAGGTTGATATGTTATCAAGGTAGATGGAATACTATGTTTACCTTAAGGA 357
QY 364 TGCAGCTCATCGGATATATTCGACAAAGAGAGATTAACCTCAGAGAGAGAACG 423
Db 358 TGCTCCCAACCGGATAACGTCCTGACAAAGAGGAAATCAATCGACAAAAACAAGAGCA 417
QY 424 CAGTCAATATCAT-----AATCAAGAGCAGATATGCTGTTGCTGACGACGAGC 474
Db 418 TAGTCAACATCGTGAAGGTGGAACTCCAAAGAACGATGCTGTTGCTTGGCACTTC 477
QY 475 CCAAGGAGTTATACACGAGATGATGGGTATATCTTCAATGATCTGATATCAATGAGGA 534
Db 478 GCAAGGAGCTATATACTACAGATGATGTTATATCTTTAATGCTTCTGATATCATAGAGGA 537
QY 535 CACGGGTGATGTTATATGCTTCCCTCAAGGCAACCATTAACATTAATCTTCCATGAGATGA 594
Db 538 TACTGTTGATGCTTATATGCTTCCCTCATGGAGATCATTAACCATTAATCTTCCATGAAGATGA 597
QY 595 GTTATCAGCTAGCAGGTTAGTCTGCTGCAAGAGCCT----- 629
Db 598 GTTATCAGCTAGCAGGTTGGCTGCTGCAAGAGCCTTCCCTATCTGCTGAGGAAATCTGTC 657
QY 630 -----ATTGGAATGGGAAGCAGGATCTCGTCTCTTCTCAAGTTCTAGTTATATATGC 681
Db 658 AAATTCAGAAACCTATCCCGACAAAATAGCATTAACACTTCAAGAAACAACTGGGTACC 717
QY 682 AAATCCAGCTCAACCAAGATTTGTAGAGAACCAAAATCTGACTGTCTCCTCACTTATCA 741
Db 718 TTCTGTAGCAATCCAGGAACCTACAAATACTAACCAAGCAACCAAGCAACACTAACAG 777
QY 742 TCA---AAATCAAGGGGAAACATTTCAAGCCTTTTCAAGTATGTTATGCTTAAACCCCT 798

778 TCAAGCAAGTCAAAGTAAATGATAGTCTCTTGAACAGCTCTACAACTGCCTTT 837
799 ATCAGAAACGCAATGGAATCTGATGGCTTATTTTCAGCCAGCGCAAAATCACAAGTCG 858
838 GAGTCAACGACATAGAAATCTGATGGCTTGTCTTTGATCCAGCAAAATCACAAGTCG 897
859 AACCGCAGAGTGTAGCTGTCCTCATGGTAAACATTAACCACTTTATCCCTTATGAACA 918
898 AACAGCTAGAGGTGTGAGTGGCCACACGCGAGATCATTTACCACTTCATCCCTTATCTCA 957
919 AATGCTGAATTTGAAAAACAAATGCTCGTATATTCCTTATTCCTTATCGTTTCAAAACCA 978
958 AATGCTGAATTTGAGAACGATCGCTCGTATATTCCTTATTCCTTATCGTTTCAAAACCA 1017
979 TTGGGTACCAAGATTAAGACAGCAACCAAGTCCCAATCGATCCGGAACCTAGTCC 1038
1018 TTGGGTACCAAGATTAAGACAGCAACCAAGTCCCAATCGATCCGGAACCTAGTCC 1077
1039 AAGTCCGCAACTGACCAAAATCCTCAACGAGCTCCAAGCAATCCAATTTGATGAAATTT 1098
1078 AGGCCCGCAACTGACCAAAATCCTCAAAATGACTCAA-----ATTCTCTTTT 1125
1099 GGTCAAAAGAGCTGTTCGAAAGTAGGCGATGGTTATGCTTTTGGAGGAATGGAGTTTC 1158
1126 GGTAGTACGCTGTGACGAAAGTTGGGAAGGATATGTTTTCGAAAGAGGGGATCTC 1185
1159 TCGTTATATCCAGCCAGGATCTTTTCAGCGAAACAGCAGCAGGCAATTTAGTACAACT 1218
1186 TCGTTATGCTTTTCGAAAGATTTACCATCTGAACCTGTAAATCTTTGAAGCAAGTT 1245
1219 GGCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAACTGACCTCCCATCTAG 1278
1246 ATCAAAACAGAGAGTGTTCACACACTTTAAGTCTTAAAGAAAGAAATTTGCTCTCTCG 1305
1279 TGATCGAGAAATTTACATAAGGCTTATGACTTACTAGCAAGAAATTCACCAAGATTTACT 1338
1306 TGACCAAGAAATTTATGATAAGCAATATACTGTTAACTGAGGCTCATAAGGCTTGT 1365
1339 TGATAATAAGGCTCGCAAGTTGATTTTTCAGGCTTTGGATTAACCTGTTGGAACGACTCA 1398
1366 TGAATAAGGCTGTAATCTGATTTCCAGGCTTAGCAAAATTTATGACGCTTGA 1425
1399 GGATGTCNCAAGTATGAAGTCAAGTTAGTGGAGATATCTTGCTTCTAGTCCGAT 1458
1426 TGATGAATCGACTAATAAGAAAGAAATTTGGTAGATGATTTATTTGGCACTTCTAGCAACCAAT 1485
1459 TCGTATCCCAAGCTTTAGGAACCAAAATCGCAAAATTTACCTACATGATGATGAGAT 1518
1486 TACCATCCAGAGGCTTGGCAACCAAAATTTCTCAAAATTTGATGATGATGATGATGATGAT 1545
1519 TCAAGTAGCCAAAGTTGGCAGCAAGTACACAAAGAGAGCGGTATATCTTTGATCTCTCG 1578
1546 TCGTATTTGCTCAATAGCTGATAGTATACAGCTCAGATGTTTATATTTTATGATGAACA 1605
1579 TGATATAACAGTATGAGGGGATGCTTATGATTAATCCACATATGACCATAGCCACTG 1638
1606 TGATATAATCAGTATGATGAGGAGATGATATGTAACGCTCATATGGGCCATAGTCACTG 1665
1639 GATTAATAAAGATAGTTTGTCTGAGCTGAGAGCGGCGAGCCAGGCTTATGCTTAAGA 1698
1666 GATTTGGAAGATAGCTTTCTGTAAGGAAAGTTTCAGCTCAAGCTCATACTACTAAAGA 1725
1699 GAAAGTTTGAACCTCTCTCGACAGACATCAGGATTCAGGAATATCTGAGGCAAAAGG 1758
1726 AAAAGGATCTTACTCTCACTTCCAGAGCGAGATGTTTAAAGCAATTCCACTGGAGATAG 1785
1759 AGCAGAGCTATCTACAAACCGCTGAAAGAGCTAAGAGGTCGCCACTTGTATGCTATGCC 1818
1786 TGCAGCAGCTATTTACATCTGTTGTAAGGGGAAAGCAAAATTCCTACTGTTGACTTCC 1845
1819 TTCAATCTCAATATATCTAGTAAGAGTCAAAACCGGTAGTTTAAATCATACCTCATATAGA 1878
1846 ATATATGTTGAGCATACAGTTGAGTTTAAAGCGTAAATTTGATTTATTTCTTCTCATAAGGA 1905

QY 1879 CCATTACCAATAACATAAAATTTGAGTGGTTTGGAGAGGCTTTATGAGGACCTTAAGG 1938
DB 1906 TCATTACCAATAATATTAATTTGTTGGTTTGTATGATCACACATACAAAGCTCCAAATGG 1965
QY 1939 GTATATCTTTGAGGATCTTTTGGCGACTGTCAAGTACTATGTGGAACATCCAAACGAACG 1998
DB 1966 CTATACCTTTGGAAGATTTGTTTGGAGCATTAAGTACTAGTAGAACACCTTGACGAACG 2025
QY 1999 TCCGATTCAGATAATGTTTGGTAACGCTAGCAGCACTGTTCAAGAAACAAAAATGG 2058
DB 2026 TCCACATTTCTAATGATGGATGGGCAATGCCATGAGCATGTGTAGGCAAGAAGACCA 2085
QY 2059 TCAAGCTGATACCAATCAACGCGAAACCAACGCGAGGAGAAACCTTCAGACAGAAAAACC 2118
DB 2086 CAGTGAAGATCAAAATTAAGAACTTCAAGCGGATGAAGAGCCAGTAGAGGAAACACCTGC 2145
QY 2119 TGAGGAGGAACCCCTCGAGAGAGA 2144
DB 2146 TGAGCGAAGATCCCTCAAGTAGAGA 2171

RESULT 9

US-09-536-784-55

; Sequence 55, Application US/09536784

; Patent No. 6573082

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 55:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2389 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 55:

US-09-536-784-55

Query Match

Best Local Similarity 43.1%; Score 987.6; DB 4; Length 2389;

Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

QY 4 TTCTATGAACCTGGTCGTCACCAAGCTGTGAGTTAAGAAAGTCTTAATCGAGTTTC 63

DB 1 TTCTTACGAGTTGGGACTGTATCAAGCTAGAACGGTTAAGGAAAA---TAATCGTGTTC 57

QY 64 TTATATAGATGGTGTATCAGGCTGGTCAAAAGGAGAGAAAATTGACACACAGATGAGTCAAG 123
Db 58 CTATATAGATGGAAAACAAGCGACGCAAAAACCGAGAAATTTGACTCTCTGTAGGGTTAG 117
QY 124 TAAAGGGGAGGGGATCAACCGCCGAACAAATNGTNATCAAGATTAACGATCAAGGTATGT 183
Db 118 CAAGCGTGAAGGATCAATGCTGAGCAAAATCGTCATCAAGNATAACAGCCAGGCTATGT 177
QY 184 GACCTCTCATGGAGACCAATATCATTAATTAATGGAAGGTTCCCTTATGATGCCATCAT 243
Db 178 CACTTCACATGGCGACCACTATCATTAATTAACATGGTAAAGTTCCTTATGACGCTATCAT 237
QY 244 CAGTGAAGAGCTCCTCATGAAAAGATCCGAATATCAGTTGAAGATTCAGACATGTCAA 303
Db 238 CAGTGAAGAATTAATCTCATGAAGAATCCAACTATAAGCTAAAGATGAGGATATTTGTAA 297
QY 304 TGAATCAAGGGTGGTTATCTCATTAAGGTAAACGGTAAATACATATCTGTTACCTTTAAGGA 363
Db 298 TGAGGTCAAGGGTGGATATGTTATCAAGGTAGATGGAAATACATATGTTTACCTTTAAGGA 357
QY 364 TGCAGCTCATGGCGATAATATTCGGACAAAAGAGAGATTAACCGTCAGAGCAGGAACG 423
Db 358 TGTGTCGCCAGCGGATAACGTCGGTACAAAAGAGGAATCAATCGACAAAACAGAGCA 417
QY 424 CAGTCATAATCAT-----AAGTCAAGAGCAGATAATGCTGTTGCTGCAGCCAGGC 474
Db 418 TAGTCAACATCGTGAAGGTGGAATCTCAAGAAACGATGGTGTCTGCTGGCTTGGCACGTTG 477
QY 475 CCAAGCAGCTTATACACAGATGATGGTATATCTTCAATGCATCTCATATCATTTAGGA 534
Db 478 GCAAGGACGCTATACACAGATGATGGTATATCTTCAATGCATCTCATATCATTTAGGA 537
QY 535 CACGGTGTATGCTTATATCTTCTCAAGCGACCAATACCATTAATTCCTTAAGAATGA 594
Db 538 TACTGTGTATGCTTATATGTTCTCATGAGATCATTAACATTAATTCCTTAAGAATGA 597
QY 595 GTTATCAGCTAGCGAGTATGCTGTCGAGAGGCT----- 629
Db 598 GTTATCAGCTAGCGAGTGGCTGCTGAGAAAGCTTCTATCTGTCGAGGAATCTGTC 657
QY 630 -----ATTGAATGGAGCAGGATCTGCTCTTCTTCAAGTCTTAGTTAATAGC 681
Db 658 AAATTCAGAAACCTATCGCGCAGCAAAATAGCGATAACACTTCAAGAACAACTGGGTACC 717
QY 682 AAATCCAGCTCAACCAAGATTTGACAGAACCAACAACTGACTGTCACTCCAACCTATCA 741
Db 718 TTCTGTGAACATCCAGGAACCTAATACTAACAAGCAACAACAGCAACCTAACAG 777
QY 742 TCA---AAATCAAGGGGAAAACATTTCAAGCCTTTTACGTGAATTTGTATGCTAAACCTTT 798
Db 778 TCAAGCAAGTCAAAAGTAATGACATTTGATAGTCTTTGAAACAGCTCTAACAACCTGCTTT 837
QY 799 ATCAGAACCCATGTGGAATCTGATGGCTTATTTTCGACCCGCGCAATCAAGTCG 858
Db 838 GAGTCAACACACATGTAGAATCTGATGGCTTGTCTTTGATCCAGCAAAAATCAAGTCG 897
QY 859 AACCGCCAGAGTGTAGTGTCTCATGTATACCATTAACCTTATCCCTTATGAACA 918
Db 898 AACAGCTAGAGGTGTGAGTGTCCACCGGAGATCAATTAACCTTATCCCTTACTCTCA 957
QY 919 AATGCTGAATTTGAAAAACGAATTTGCTGATATATTTCCCTTCTGTTATCGTTCAAAACCA 978
Db 958 AATGCTGAATTTGGAAGACGATCGCTGATATATTTCCCTTCTGTTATCGTTCAAAACCA 1017
QY 979 TTGGGTACAGATTTCAAGACCAAGCAACCAAGTCCCAATTCGACTCCGGAACCTAGTCC 1038
Db 1018 TTGGGTACAGATTTCAAGGCCAGAACAAACCAAGTCCCAACCGACTCCGGAACCTAGTCC 1077
QY 1039 AAGTCCGCAACCTGCAACCAATCTCAACAGCTCAAGCAATCCAAATTTGATGAGAAAT 1098
Db 1078 AGGCCCGCAACCTGCAACCAATCTTAAATATAGACTCAA-----ATTCTTCTTT 1125
QY 1099 GGTCAAGAGAGCTGTTCCGAAAAGTAGCGCATGTTATGTTTGGAGAGAAATCGAGTTTC 1158

RESULT 10

Db 1126 GGTATAGTCAAGTGGTACGAAAAATTTGGGGAAGGATATGTATTCGAAAAAGGCGCATCTC 1185
QY 1159 TCGTTATATCCAGCAAGGATCTTTTACAGAGAAACAGCAGCAGGCAATTCATAGCAAACT 1218
Db 1186 TCGTTATGCTTTTGGAAAGATTTACCATCTGAATCTGTAAATACTTTGAAAGCAAGTT 1245
QY 1219 GSCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAAACAGCTCCCATCTAG 1278
Db 1246 ATCAAAAACAAGAGAGTGTTCACACACTTTAACTGCTAAAAAAGAAAAATGTTGCTCTCG 1305
QY 1279 TGATCGAGAAATTTTACATAAGCTTTATGACTTTA CTAGCAAGAAATTCACCAAGATTTACT 1338
Db 1306 TGACCAAGAAATTTTATGATAAAGCATATAATCTGTAACTGAGGCTCAVAAAAGCCTTGT 1365
QY 1339 TGATAATAAAGGTCGACAAGTTGATTTTGGAGCTTTGGATTAACCTGTTGAAACGACTCAA 1398
Db 1366 TGNAAATAAGGTCGTAATCTGATTTTCAAGCTTTAGACAAATTTATAGAACGCTTGA 1425
QY 1399 GGATGTCTCAAGTGAATAAGTCAAGTTAGTGGAGATATTTTGGCTTCTTAGCTCCGAT 1458
Db 1426 TGATGAATCGACTAATAAGAAATAATTTGATAGTATGTTTTATTTGGCAATTTCTAGACCAAT 1485
QY 1459 TCGTCAATCCAGAACTTTTAGGAAAACCAATCGCAAAATTTACCTACATGATGATGAT 1518
Db 1486 TACCCATCCAGAGGACTTTGGCAAAACCAATTTCTCAAAATTTAGTATACCTGAAGCAAGT 1545
QY 1519 TCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGAGCGTTATATCTTTGATCTCG 1578
Db 1546 TCGTATTCCTCAATTTAGCTGATAAGTATATAACAAGCTCAGATGGTTACATTTTGTATGAACA 1605
QY 1579 TGATAAACCAGTGTAGGGGATGCGCTATCTTAATCCACATATGACCCATAGCAGCACTG 1638
Db 1606 TGATATAATCAGTGAAGAGAGTGCATATGTAACGCTCATATGGGCCATAGTCACTG 1665
QY 1639 GATTAAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGAGCCAGGCTTATCTTAAGA 1698
Db 1666 GATTGAAAAGATAGCTTTCTGATAAGGAAAAGTTGCAGCTCAAGCCTATATAAGA 1725
QY 1699 GAAAGTTTGAACCTCTCTCGACAGACATCAGGATTCAGGAAATPATGAGGCAAAAGG 1758
Db 1726 AAAAGGTATCTACCTCCATCTCCAGACGAGATGTTAAAGCAAAATCCAACTGAGATAG 1785
QY 1759 AGCAGAGCTATCTACAAACCGCTCAAAAGCAGCTAAGAGGTGCCACTTTGATCGTATGCC 1818
Db 1786 TGCAGCAGCTATTTACAATCGTGTGAAAGGGGAAAAACGAATTCACACTGTTGACTTCC 1845
QY 1819 TTCAATCTTCAATATACTGTAGAAGTCAAAAACCGTAGTTTAAATCATACCTCATTAATGA 1878
Db 1846 ATATATGTTGAGCATACAGTTGAGTTTAAAAACGTAATTTGATTTATTTCTCTATAAGA 1905
QY 1879 CCATTAACATAACATCAAAATTTGAGTGGTTTGAAGAGCCTTTATGAGGACCTTAAGGG 1938
Db 1906 TCATTAACATAATATTAATTTGCTTGGTTTGTATGATCACAATCAAAAGCTTCAAAATGG 1965
QY 1939 GTATACCTTTGAGGATCTTTTGGGCACTGTCAAGTACTATCTGCAACATCCAAACGAAAG 1998
Db 1966 CTATACCTTTGGAAGATTTGTTTGGGACGATTAAGTACTACGTTAGAACACCTTGAAGAACG 2025
QY 1999 TCCCATCTCAGATAATGTTTGGTGAAGCTAGCGACCATGTTTCAAGAAACAAAAAATGG 2058
Db 2026 TCCCATCTTAAATGATGGATGGGCAATGCCAGTGAAGCATGTTTGGCAAGAAAGACCA 2085
QY 2059 TCAAGCTGATACCAATCAAAACGAAAAACCAAGCGAGGAGAAACCTTCAGACAGAAAAAC 2118
Db 2086 CAGTGAAGATCCAAATAAGAACTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACACCTTGC 2145
QY 2119 TGAGGAAGAAACCCCTCGAGAAAGAGA 2144
Db 2146 TGAGCCAGAAAGTCCCTCAAGTAGAGA 2171

US-08-961-527-355
; Sequence 355, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-355

Query Match 34.7%; Score 794.8; DB 4; Length 973;
Best Local Similarity 98.7%; Pred. No. 8.1e-234;
Matches 810; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
QY 1 TTGTTCTTATGAATGGTGGTCAACCAAGCTGTGAGGTTAAGAAAGAGTCTAATCGAGT 60
DB 154 TGTTCCTTAGAGCTTGGACGTTACCAAGCTGTGAGGTTAAGAAAGAGTCTAATCGAGT 213
QY 61 TTCTTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAACTTGACACCCAGATGAAGT 120
DB 214 TGCTTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAACTTGACACCCAGATGAAGT 273
QY 121 CAGTAAGAGGGAGGGGATCAAGCCGAAACAAATGNTNATCAAGATTACGATCAAGGTTA 180
DB 274 CAGTAAGAGGGAGGGGATCAAGCCGAAACAAATGNTNATCAAGATTACGATCAAGGTTA 333
QY 181 TGTGACCTCTCATGAGACCATTCATTACTATATGCGATTAATGCGAGTTCTTATGATGCCAT 240
DB 334 TGTGACCTCTCATGAGACCATTCATTACTATATGCGATTAATGCGAGTTCTTATGATGCCAT 393
QY 241 CATCAGTGAAGAGCTCTCATGAAGATCCGATTAATCAGTTTGAAGATTTCAGACATTTGT 300
DB 394 CATCAGTGAAGAGCTCTCATGAAGATCCGATTAATCAGTTTGAAGATTTCAGACATTTGT 453
QY 301 CAATGAATCAAGGGTGGTTATGTCTATTAAAGGTAACCGGTAATACTATGTTNACCTTAA 360
DB 454 CAATGAATCAAGGGTGGTTATGTCTATTAAAGGTAACCGGTAATACTATGTTNACCTTAA 513
QY 361 GGATGAGCTCATGGGATTAATATTCGGGCAAAAGAGAGATTAAACGTCAGAGCAGGA 420
DB 514 GGATGAGCTCATGGGATTAATATTCGGGCAAAAGAGAGATTAAACGTCAGAGCAGGA 573

QY 421 AGCAGTCTATATCATTAACCTCAAGAGCAGATATGCTGCTGCTGAGGAGCCCAAGG 480
DB 574 AGCAGTCTATATCATTAACCTCAAGAGCAGATATGCTGCTGCTGAGGAGCCCAAGG 633
QY 481 ACCTTATACAAACCGATGATGGGTATATCTTCAATGATCATCTGATATCATTTGAGGACACGGG 540
DB 634 ACCTTATACAAACCGATGATGGGTATATCTTCAATGATCATCTGATATCATTTGAGGACACGGG 693
QY 541 TGATGCTTATATCTGCTTCTCACGGGACCAATTACCATTAATCTCTTAAGAAATGATGTTATC 600
DB 694 TGATGCTTATATCTGCTTCTCACGGGACCAATTACCATTAATCTCTTAAGAAATGATGTTATC 753
QY 601 AGCTAGGAGTGTAGCTGCTGCTGAGAGGCTTATGGAATGGGAAGCAGGATCTGCTCTTC 660
DB 754 AGCTAGGAGTGTAGCTGCTGCTGAGAGGCTTATGGAATGGGAAGCAGGATCTGCTCTTC 813
QY 661 TTCAAGTTCTAGTATTAATTAATCAAGTCCAGTCAACCAAGATTGTGAGAAACCAATCT 720
DB 814 TTCAAGTTCTAGTATTAATTAATCAAGTCCAGTCAACCAAGATTGTGAGAAACCAATCT 873
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 780
DB 874 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 933
QY 781 ATTGTATGCTAAACCTTATCATCAAGCAGCATGTGGAATCTG 821
DB 934 ATTGTATGCT-AAACCTTATCATCAAGCAGCATGTGGAATCTG 973

RESULT 11
US-08-961-527-258
; Sequence 258, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 258:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-258
Query Match 29.6%; Score 677; DB 4; Length 1684;
Best Local Similarity 99.9%; Pred. No. 1.9e-197;

Matches 688; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1602 ATCCCTATGTAACCTCCATATGACCATAGCCATGCGATTAATAAGATAGATTGCTG 1661
Db |||||
1 ATCCCTATGTAACCTCCATATGACCATAGCCATGCGATTAATAAGATAGATTGCTG 60
QY 1662 AGCTGAGAGAGGGGCGCCAGGCTTATGCTAAAGAAAGTTTGACCCCTCCTTCGA 1721
Db |||||
61 AAGCTGAGAGAGGGGCGCA-CGAGGCTTATGCTAAAGAAAGTTTGACCCCTCCTTCGA 119
QY 1722 CAGACCATCAGGATTCAGGAAATCTAGGCAAAAGGAGGAGGAGCTTCTACACCGCG 1781
Db |||||
120 CAGACCATCAGGATTCAGGAAATCTAGGCAAAAGGAGGAGGAGCTTCTACACCGCG 179
QY 1782 TGAAGCAGCTAAGAGGTCCTTATGATCGTATGCTTCAATCTTCAATATATCTGTAG 1841
Db |||||
180 TGAAGCAGCTAAGAGGTCCTTATGATCGTATGCTTCAATCTTCAATATATCTGTAG 239
QY 1842 AGTCAAAAACGGTAGTTTAATCATACCTCATTTAGACCAATTAACCAATCAAAATTTG 1901
Db |||||
240 AGTCAAAAACGGTAGTTTAATCATACCTCATTTAGACCAATTAACCAATCAAAATTTG 299
QY 1902 AGTGGTTTACGAGGCTTATGAGGACCTTAAGGGTATATCTTTGAGGATCTTTGG 1961
Db |||||
300 AGTGGTTTACGAGGCTTATGAGGACCTTAAGGGTATATCTTTGAGGATCTTTGG 359
QY 1962 CGACTCTCAAGTACTATGTGCAATCCAAACGAGCTCGCATTCAGATAATGCTTTTG 2021
Db |||||
360 CGACTCTCAAGTACTATGTGCAATCCAAACGAGCTCGCATTCAGATAATGCTTTTG 419
QY 2022 GTAACGCTGACGACCTTGTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 2081
Db |||||
420 GTAACGCTGACGACCTTGTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 479
QY 2082 AAAACCAAGCAGGAGGAACTTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 2141
Db |||||
480 AAAACCAAGCAGGAGGAACTTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 539
QY 2142 AGAAGCGGAAAGCAGGAACTTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 2201
Db |||||
540 AGAAGCGGAAAGCAGGAACTTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 599
QY 2202 CAGAGGAATCAGAGAACTTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 2261
Db |||||
600 CAGAGGAATCAGAGAACTTCAAGAAACAAATAATGCTCAAGCTGTACCAATCAACGG 659
QY 2262 CTGAAGATTCTTGAAGAAATCCAGAT 2290
Db |||||
660 CTGAAGATTCTTGAAGAAATCCAGAT 688

RESULT 12

US-08-961-083-181
; Sequence 181, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 181:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-181
Query Match 17.5%; Score 401; DB 3; Length 1342;
Best Local Similarity 65.7%; Pred. No. 1.1e-112;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;
QY 23 CACCAAGCTGCTCAGGTTAAAGAAAGAGTCTAATCGAGTTTCTTATATAGATGCTGATCAG 82
Db |||||
8 CAGCATCGTTCGACGAGAAATAAGGACAATAATCGTGTCTCTTATGTGGATGCGACCGAG 67
QY 83 GCTGCTCAAAAGCGCAGAAAACTTGACACAGATGAAGTCAAGTAAAGAGGGAGGGGATCAAC 142
Db |||||
68 TCAAGTCAAGAAAGTGAACCTTGACCCAGACCCAGGTTAGCCAGAAAGAAAGAAATTCAG 127
QY 143 GCCGAACAATNGTNTCAAGATTACGGATTCAAGGTTATGTGACCTCTCATGGAGACCAT 202
Db |||||
128 GCTGAGCAATTTGTAATCAAAATTAAGATCAGGCTGTATGACGTCACCGGTGACCCAC 187
QY 203 TATCATTTACTAATGCGCAAGGTTCCCTTATGATGCCATCATCAGTGAAGAGTCCCTCATG 262
Db |||||
188 TATCATTTACTAATGGGAAAGTTCCTTATGATGCCCTCTTTAGTGAAGAACTCTTGATG 247
QY 263 AAAGATCCGAATTTACGTTGAAGGATTGACATTTGCAATGAAATCAAGGGTGGTTAT 322
Db |||||
248 AAGGATCCAAATCATCACTTAAAGACGCTGATATTTGCAATGAAGTCAAGGGTGGTTAT 307
QY 323 GTCATTAAAGTAAAGGTTAAATCTATGTTACCTTTAAGGATGACGCTCATGCGGATAT 382
Db |||||
308 ATCATCAAGGTCGATGGAAATATTTATGTTCTACCTGAAAGATGCGAGCTCATGCTGATA 367
QY 383 ATTCGGACAAAGAGAGATTAAAGCTCAGAGCAGGAAACGAGTCATAATCATAACTCA 442
Db |||||
368 GTTCGAATTAAGATGAATCAATGCTCAAAACCAAGAAC--ATGCAAGATATAGAG 424
QY 443 AGAGCAGATAATGCTGTTGCTGACGCCAGAGCCCAAGGACGTTTATACACGGATGATGGG 502
Db |||||
425 AAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGAGCAGATATACGCAAAATGATGGT 484
QY 503 TATATCTTCAATGCAATCTGATATCAATGAGGACAGGAGTATGCTTATATCGTCTCTCAC 562
Db |||||
485 TATGTCCTTTAATCCAGCTGATATTTATCGAAGATACGGGTAATGCTTATATGCTCTCAT 544
QY 563 GCGACACCATTAACCATTAATCTTCAAGATGATGTTATCAGCTAGCGAGTGTAGCTGTGCA 622
Db |||||
545 GGAGGTCACTATCACTACATTTCCCAAAAGCGATTTATCTGCHAGTGAATTAGCAGAGCT 604
QY 623 GAAGCTTATGGAATGGGAAGCAGGGATCTCGTCTCTTCTTCAAGTTTCTAGTTTATATGCA 682
Db |||||
605 AAAGCACAATCTGGCTGGAAATAATGCAACCGAGTCAGTTAAGCTATTCTTCAACAGCT 664
QY 683 AATCCAGCTCAACCAAGATGTCAGAGAACCAATCTGACGTGTGCTCACTCAACTTATCAT 742
Db |||||
665 AGTGACATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 712
QY 743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTTGATGCTTAAACCCCTTATCA 802

Db 713 GCATTAATCTGAAATCTCCAGAGTCTTTTGAAGGAACCTATGATTCACCTAGCGCC 772
QY 803 GAAGCCATGTGGAATCTGATGGCTTATTTTCGACCCAGCGCAATCACAAGTCGAACC 862
Db 773 CAAGCTTACAGTGAATCAGATGGCTGTCTTTGACCTCTGATGATTTATGATCGTACA 832
QY 863 GCCAGAGGTGTAGCTGCTCCCTCATGGTAACCATACCACTTTATCCCTTATGAACAAATG 922
Db 833 CCAATGGAGTGTGATTCGCGATGGCGACCATACCACTTTATTCCTTACAGCAAGCTT 892
QY 923 TCTGAATGGAAACGAATTCGTCGTATTTATTCCTTCGT 964
Db 893 TCTGCCTTAGAAGAAAGATTGCCAGATGGTGCCTATCAGT 934

RESULT 13

US-09-536-784-181
; Sequence 181, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB34023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 181:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 181:
US-09-536-784-181

Query Match 17.5%; Score 401; DB 4; Length 1342;
Best Local Similarity 65.7%; Pred. No. 1.1e-112;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;
QY 23 CACCAAGTGTGTGAGTTAAGAAAGTCTTAATCAGTTCTTATATAGATGGTATCAG 82
Db 8 CAGCATCTTCGCGAGAAATTAAGACAATATCGTCTCTTATGTGATGGCAGCCAG 67
QY 83 GCTGTCAAAAGCGAGAAACTTGACACAGATGAAGTCAGTAAGAGGGAGGGATCAAC 142
Db 68 TCAAGTCAGAAAGTGAACCTTGACACAGCAGGTTAGCCAGAAAGAAAGAAATTCAG 127
QY 143 GCCGAAACAAATNGTGNATCAAGATTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT 202

Db 128 GGTGAGCAAAATGTATCAAAATTAAGATCAGGGCTATGTAACGTCAAGGGTGACAC 187
QY 203 TATCATTAATAATGCGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCCTCATG 262
Db 188 TATCATTAATAATGCGAAAGTTCTTATGATGCCCTCTTATGTAAGAACTCTTGATG 247
QY 263 AAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTGCTCAATGAATCAAGGGTGGTTAT 322
Db 248 AAGGATCCAAACTATCAACTTAAAGACGCTGATATTGTCAATGAAGTCAAGGGTGGTTAT 307
QY 323 GTCATTAAGGTAACCGTAAATTAATTAATGCTTAAAGATGCAGCTCAAGGGATTAAT 382
Db 308 ATCATCAAGGTCGATGGAATTAATTAATGCTTAAAGATGCAGCTCAAGGGATTAAT 367
QY 383 ATTGCGACAAAAGAGAGATTAAACGTCAGAACGAGCAAGCGAGTCAATATCAATACTCA 442
Db 368 GTTCGAACTAAGATGAATCAATCGTCAAAACAAGAAC---ATGTCAAAGATAATGAG 424
QY 443 AGAGCAGATAATGCTGTTGCTGCGACGAGCCAAAGGAGCTTATACAAGGATGATGGG 502
Db 425 AAGGTTAACTCTAAATGTTGCTGTAGCAAGGCTCTCAGGACGATATACGACAAATGATGGT 484
QY 503 TATATCTTCAATGCATCTGATATCATTCAGGACGAGGCTGATGCTTATATCGTTCCTCAC 562
Db 485 TATGCTTTTAAATCCAGCTGATATTCGAAGATACGCGGTAATGCTTATATCGTTCCTCAT 544
QY 563 GCGAGCACTTACCATTACATTCCTAAGATGAGTATCAGCTAGCGAGTTAGCTGTGCA 622
Db 545 GGAGGTCATCTACTACATTCCTCAAAAGGATTTATCTGCTAGTGAATTAGCAGAGCT 604
QY 623 GAAGCCTATTGGAATGGGAAGAGGATCTGTCCTTCTTCAAGTTCTAGTTATAATGCA 682
Db 605 AAAGCACATCTGGCTGGAATAATATGCAACCGAGTCAGTTAAGCTATTCTTCAACAGCT 664
QY 683 AATCAGCTCAACCAAGATTGTGAGAGAACCAACAACTGACTGTCTCATCCAACTTATCAT 742
Db 665 AGTGACAATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 712
QY 743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTAATGTATGCTTAAACCTTATCA 802
Db 713 GCAATTAATCTGAHAATCTCCAGATCTTTGAAGAACTCTATGATTCACCTAGCGCC 772
QY 803 GAACGCCATGTGGAATCTGATGGCTTATTTTCGACCCAGCGCAAAATCACAAGTCAAC 862
Db 773 CAAGGTTACAGTGAATCAGATGCGCTGGTCTTTGACCCCTGCTAAGATTATCAGTCGTACA 832
QY 863 GCAGAGGTGTAGCTGTCCCTCATGTATACCATACACATTTATCCCTTATGAACAAATG 922
Db 833 CCAATGGAGTTGCGATTCGATTCGATGCGACCATTAACCACTTTATTCCTTACAGCAAGCTT 892
QY 923 TCTGAATTCGAAACGAATTCCTGCTATTTATTCCTTCCTTCGT 964
Db 893 TCTGCCTTAGAAGAAAGATTGCCAGATGGTGCCTATCAGT 934

RESULT 14

US-09-468-656A-7
; Sequence 7, Application US/09468656A
; Patent No. 6582706
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-444
; CURRENT APPLICATION NUMBER: US/09/468,656A
; CURRENT FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7


```

; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-09-468-656A-7

Query Match      17.5%; Score 401; DB 4; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.2e-112;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

20 23 CACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTTCTTATATAGATGGTGATCAG 82
    |||
79 CAGCATCGTCGAGGAAATATAGCAATATCTGTCTCTTATGTGATGGAGCCAG 138
    |||
83 GCTGGTCAAAAGGAGGAAACTTGCACCAAGATGAGTCAAGTACAGAGGGGAGGATCAAC 142
    |||
139 TCAAGTCAGAAAGAGTGAAGAACTTGACACCAAGCCAGGTTAGCCAGAAAGAAATTCAG 198
    |||
143 GCCGACCAATNGTNATCAAGNTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT 202
    |||
199 GCTGAGCAAAATGTAATCAAAATTTACAGATCAGGGCTATGTAACTGACAGGTCACCC 258
    |||
203 TATCATTTACTATATATGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCTCATG 262
    |||
259 TATCATTTACTATATGGAAGTTCTTATGATGCCCTCTTTAGTGAAGAACTCTTGATG 318
    |||
263 AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGTCATGAAATCAAGGGTGGTTAT 322
    |||
319 AAGGATCCAACTATCAACTTAAAGCGCTGATATTTGTCATGAAGTCAAGGGTGGTTAT 378
    |||
323 GTCATTAAGGTAACGGTAAATCTATGTGTTACTTAAAGATGAGTCAAGTCTCATGGGATAAT 382
    |||
379 ATCATCAAGTCGATGGAATATTTATGTCTACCTGAAGATGAGTCAAGTCTCATGCTGATAAT 438
    |||
383 ATTCGACAAAGAGATTAACGTCAAGAGCAGGAAAGCAGTCAATATCAATCAACTCA 442
    |||
439 GTTCGAAGTAAAGATGAATCAATCGTCAAAAGCAAGAAC---ATGTCAAAGATTAATGAG 495
    |||
443 AGAGCAGATTAATCTGTGTCGACGAGGAGCCAGGATTTATACAAAGGATGATGGG 502
    |||
496 AAGTTAACTCTAATGTGTCGACGAGGATCTCAGGAGCATATACGACAAATGATGGT 555
    |||
503 TATATCTTCAATGCATCTGATATCAATTTAGGACACAGGGTGATGTTATATCGTTCCTCAC 562
    |||
556 TATGTCCTTTAATCCAGTGTATATATCGAAGATACGCGGTAATGCTTATATCGTTCCTCAT 615
    |||
563 GCGGACCATTACCATTTACATCTCTAGAAATGAGTTATCAGTACGAGTATGCTGCTGCA 622
    |||
616 GGAGGTCATCTACATACATCTCCAAAGGAGTATTTATCTGCTAGTGAATTAGCAGCAGCT 675
    |||
623 GAAGCCTATTGGAATGGGAGCAGGAGTCTCGTCTCTTCAAGTCTGACTCTCAACCTTATCAT 682
    |||
676 AAGACCATCTGCTGGGAAATAATGCAACCGAGTCAGTTAGTACTATCTTCAACAGCT 735
    |||
683 AATCCAGCTCAACCAAGATTTGTGAGAGAACCAAACTGACTGTCTCACTCCAACTTATCAT 742
    |||
736 AGTGACAATAACACGCAATCTGTAGCAA-----AGGATCAACTGCAAGCCA 783
    |||
743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTAATTTGATGCTAAACCCCTTATCA 802
    |||
784 GCAAAATAATCTGAAATCTCCAGAGTCTTTTGAAGGAATCTCTATGATTCACCTAGCGCC 843
    |||
803 GAAGCCCATGTGAATCTGATGSCCTTATTTTCGACCCAGCGCAATCACAGTCCGAC 862
    |||
844 CACAGTTACAGTGAATCAGATGSCCTGGTCTTTGACCCCTGCTAAGATTAATCAGTCGTA 903
    |||
863 GCCAGAGGTGATGCTGCTCCCTCATGTGTAACCATTAACCATTTTATCCCTTATGAACAAATG 922
    |||
904 CCAATTTGAGGTTCGATTCGATGCGATGGGACCATTAACATTTATTCCTTATCCTTACAGCAAGCTT 963
    |||
923 TCTGAATGGAAACGAAATTTGCTCGTATTTATTTCCCTTCGT 964
    |||
964 TCTGCCTTAGAAGAAAGATTTGCGAATGCTGCTATCAGT 1005
    |||

RESULT 15
US-08-961-527-192/c
; Sequence 192, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-192

Query Match      17.5%; Score 401; DB 4; Length 6867;
Best Local Similarity 65.7%; Pred. No. 3.4e-112;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

Qy 23 CACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTTCTTATATAGATGGTGATCAG 82
    |||
Db 6707 CAGCATCGTTCGAGGAAATATAGGACATATATCGTGTCTCTTATGTGATGGAGCCAG 6648
    |||
Qy 83 GCTGGTCAAAAGGAGGAAACTTGTGACAGATGAAGTCAGTAAGAGGGAGGGATCAAC 142
    |||
Db 6647 TCAAGTCAGAAAGTGAAAACTTGACACAGACCAGGTTAGCCAGAAAGAAATTCAG 6588
    |||
Qy 143 GCCGACCAATNGTNATCAAGNTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT 202
    |||
Db 6587 GCTGAGCAAAATGTAATCAAAATTTACAGATCAGGGCTATGTAACTGATCAACCGTGA 6528
    |||
Qy 203 TATCATTTACTATATGCGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCCCTCATG 262
    |||
Db 6527 TATCATTTACTATATGCGGAAAGTTCTTATGATGCCCTCTTTAGTGAAGAACTTTGATG 6468
    |||
Qy 263 AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGTCAATGAAATCAAGGGTGGTTAT 322
    |||
Db 6467 AAGGATCCAAACTATCAACTTAAAGACGCTGATTTGTCAATGAAGTCAAGGGTGGTTAT 6408
    |||
Qy 323 GTCATTAAGGTAACGGTAAATATCTATGTTNATACCTTAAAGATGCGAGTCTCATGGGATAAT 382
    |||
Db 6407 ATCATCAAGGTCGATGGAATAATTTATGTCTACCTGAAAGATGCGAGTCTCATGCTGATAAT 6348
    |||
Qy 383 ATTCGACAAAGAGATTAACGTCAAGAGCAGGAGCAGTCAATATCAATCAACTCA 442
    |||
```

Db 6347 GTTCGAACAAAGATGAATCAATCGTCAAAACACAGAAC---ATGTCAGAGATTAATGAG 6291
Qy 443 AGACGAGATATGCTGTTGCTGACGACAGACCCAGGACGTTATACACGATGATGG 502
Db 6290 AAGGTTAACTCTAATGTTGCTGACGAGGCTCAGGAGCATATACGACAAATGATGT 6231
Qy 503 TATATCTTAAATGATCATGATATCATTTGAGACACGGGTGATGCTTATATGTTCTCTAC 562
Db 6230 TATGCTTTTATCCAGCTGATTAATTCGAAGATACGGGTATGCTTATATGTTCTCTCAT 6171
Qy 563 GGCACCATTAACATTAATTCCTAAGATCAGTTATCAGTATCAGCTAGGCTTACGCTGCA 622
Db 6170 GGAGGTCATATCACTACATTCACAAAGCGATTATCTGCTAGTGAATAGCAGCAGCT 6111
Qy 623 GAAGCTATTGGAATGGAGACGGGATCTGCTCTTCTTCAAGTTCTAGTTTATATGCA 682
Db 6110 AAGCACATCTGGCTGGAAAAAATATGCAACGGAGTCAGTTAAGCTATCTTCAACAGCT 6051
Qy 683 AATCCAGCTCAACCAAGATGTCAGAGAACCAATCTGACTGTCACTCCCACTTATCAT 742
Db 6050 AGTGACAAATACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 6003
Qy 743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATGTTATGCTTAAACCTTATCA 802
Db 6002 GCAATTAATCTGAATCTCCAGATCTTTTGAAGGAATCTTATGATTCACCTAGCGCC 5943
Qy 803 GAAGCCATGTGGAATCTGATGGCTTATTTTGAACCCAGGCAATCACAAGTCGAACC 862
Db 5942 CAAGTTACAGTGAATCAGATGGCTGTCTTTGACCTGTAGATTTATCAGTGTACA 5883
Qy 863 GCACAGGTGATGTCCTTCATGTTGTAACCATTAACCACTTATCCCTTATGAACAAATG 922
Db 5882 CCAATAGGATGCGATTCCGATGGCGACCATTACCACCTTATTCCTTACAGCAAGCTT 5823
Qy 923 TCTCAATTTGAAAAAGAACTGCTGCTGATTTATTCCTTCTG 964
Db 5822 TCTGCTTAGAAGAAAGATTGCGAGATGTTGCTATCAT 5781

RESULT 16
US-08-743-637B-34/c
; Sequence 34, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:

; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 841 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; US-08-743-637B-34

Query Match 11.9%; Score 272.4; DB 2; Length 841;
Best Local Similarity 92.9%; Pred. No. 2.7e-73;
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;

Qy 1956 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCGCATTCAGATAATG 2015
Db 841 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCGCATTCAGATAATG 783
Qy 2016 GTTTTGGTAACGCTAGCGACCATGTTCAAAGAAACAAATAATGGTCAAGCTGATACCAATC 2075
Db 782 GTTTTGGTAACGCTAGCGACCATGTTCAAAGAAACAAATAATGGTCAAGCTGATACCAATC 723
Qy 2076 AAACGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAAACCTGAGGAGAAACCCCTC 2135
Db 722 AAACGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAAACCTGAGGAGAAACCCCTC 663
Qy 2136 GAGAAGAGAAACCGAAACGAGGAGAAACCGAGGTCCTCCAAACCAACAGAGGA----- 2187
Db 662 GAGAAGAGAAACCGAAACGAGGAGAAACCGAGGTCCTCCAAACCAACAGAGGA----- 603
Qy 2188 -----ACCAGAAGAAATCACCAGAGGAATCAGAGAACTCAGGTCGAGACTGAAAGG 2240
Db 602 AAGATCACCAGAGAAATCACCAGAGGAATCAGAGAACTCAGGTCGAGACTGAAAGG 543
Qy 2241 TTGAAGAAAACTGAGAGAGGCTGAAAGATTTACTTGGAAAAATCCAGGAT 2290
Db 542 T---AAGAAAACTGAGAGAGGCTGAAAGATTTACTTGGAAAAATCCAGAA 496

RESULT 17

US-08-526-840B-34/c
; Sequence 34, Application US/08526840B
; Patent No. 6001564
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
; NUMBER OF SEQUENCES: 177
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,840B

FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/304,732
FILING DATE: 12-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586.90012
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 841 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
US-08-526-840B-34

Query Match 11.9%; Score 272.4; DB 3; Length 841;
Best Local Similarity 92.9%; Pred. No. 2.7e-73;
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;

QY 1956 TTTTGGCGACTGCAAGTACTATGTGCAATCAACGAAAGTCCGCAATTCAGATAATG 2015
DB 841 TTTTGGCGACTGCAAGTACTATGTGCAATCAACGAAAGTCCGCAATTCAGATAATG 783

QY 2016 GTTTTGTGAACGTACGACCATGTTCAAGAGAAACAAATGTTCAAGCTGATACCAATC 2075
DB 782 GTTTTGTGAACGTACGACCATGTTCAAGAGAAACAAATGTTCAAGCTGATACCAATC 723

QY 2076 AAACGGAAAAACCAACGAGGAGAAACCTCAGACAGAAAAACCTTGAGGAAGAAACCCCTC 2135
DB 722 AAACGGAAAAACCAACGAGGAGAAACCTCAGACAGAAAAACCTTGAGGAAGAAACCCCTC 663

QY 2136 GAGAGAGAAACCGGAAAGGAGAAACAGAGTCTCCAAAAACCAACAGAGGA 2187
DB 662 GAGAGAGAAACCGGAAAGGAGAAACAGAGTCTCCAAAAACCAACAGAGGAAG 603

QY 2188 -----ACCAGAAGATCACAGAGGAATCAGAGAAACCTCAGGTCGAGACTGAAAGG 2240
DB 602 AGAATCACCAAGAAATCACAGAGGAATCAGAGAAACCTCAGGTCGAGACTGAAAGG 543

QY 2241 TTGAGAAAAACCTGAGAGGCTGAAGATTACTTTGAAAAATCCAGCAT 2290
DB 542 T---AAGAACTGAGAGGCTGAAGATTACTTTGAAAAATCCAGAT 496

RESULT 18
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 2.8%; Score 64.4; DB 1; Length 7218;
Best Local Similarity 3.9%; Pred. No. 1e-08;
Matches 11; Conservative 180; Mismatches 91; Indels 0; Gaps 0;

QY 2008 AGATAATGTTTGGTAACTAGCGTAGGACCATGTTCAAGAAACAAATGGTCAAGCTGA 2067
DB 1453 AGATAAGAAATTTGTACTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1394

QY 2068 TACCAATCAACGGAACCAAGCGAGAGAAACCTCAGACAGAAAAACCTGAGGAAGA 2127
DB 1393 RRR 1334

QY 2128 AACCCCTCAGAGAGAAACCGCAAGCGAGAGAAACAGAGTCTCCAAAAACCAACAGAGA 2187
DB 1333 RRR 1274

QY 2188 ACCAGAGAATCACAGAGGAATCAGAGAAACCTCAGGTCGAGACTGAAAGTTGAAGA 2247
DB 1273 RRR 1214

QY 2248 AAAACTGAGAGGCTGAAGATTACTTTGAAAAATCCAGGA 2289
DB 1213 RRR 1172

RESULT 19
US-08-257-073-4
Sequence 4, Application US/08257073
Patent No. 5766597
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
APPLICANT: de Taisne, Charles
APPLICANT: Tine, John A.
TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue, 25th Floor
CITY: New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 454310-2570
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CUFTMS
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2223 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-257-073-4
Query Match 2.5%; Score 56.2; DB 1; Length 2223;
Best Local Similarity 48.0%; Pred. No. 1.5e-06;
Matches 160; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1957 TTGGGCACTGTCAGTACTATGTCGAACATCCAAACGACGTCGCGATTCAGATAATGG 2016
Db 1863 TTATCAAAATGGAACCAAAAATATTTCAACAAAATATTTCTAGAAAATGA 1922
QY 2017 TTTTGGTAACGCTAGCGACCATGTTCAAGAACCAAAAATGTCAGCTGTACCAATCA 2076
Db 1923 TGTTCTTAATCAAGAACGGAGGAAGAAATGGAATAACAGTTGAAGCAATCACCAGCA 1982
QY 2077 AACGGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAAACCTGAGGAGAAACCCCTCG 2136
Db 1983 AATAGAAGCTGAAGTGGATGCCCTGCACCAAAAATTAAGGAGAGAGAAAAGAAAA 2042
QY 2137 AGAAGAAACCGCAAGCGAGAACCGAGTCTTCCAAAACCAACAGAGGAAACCGAGAGA 2196
Db 2043 AGAAAAAGAAAAGGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAGA 2102
QY 2197 ATCAACAGAGGAATCAGAGAACCTCAGGTGAGACTGAAAGGTTGAAGAAAAACTGAG 2256
Db 2103 AAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAGA 2162
QY 2257 AGAGGCTGAAGATTACTTTGAAAAATCCAGGA 2289
Db 2163 ACAGAGAGAGAGAGAGAGAGAAATAGTACCAGA 2195

```

```

RESULT 20
US-09-640-419C-6
; Sequence 6, Application US/09640419C
; Patent No. 6630615
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L
; APPLICANT: Crasta, Oswald R
; APPLICANT: Hu, Xu
; APPLICANT: Lu, Guihua
; TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
; FILE REFERENCE: 35718/199009 (5718-92)
; CURRENT APPLICATION NUMBER: US/09/640,419C
; CURRENT FILING DATE: 2000-08-17

```

```

; PRIOR APPLICATION NUMBER: 60/149,656
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/206,405
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(519)
;
US-09-640-419C-6
Query Match 2.2%; Score 51.4; DB 4; Length 865;
Best Local Similarity 50.2%; Pred. No. 2.3e-05;
Matches 127; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 2002 GCATTCAGATAATGTTTGGTTAACGCTAGCGACCATGTTCAAAGAAACAAAATGTC 2061
Db 30 GCATGATGAGAAATTTAGGAGGAATCGAATGTCGAATCGGAGAGCACAATAAGAGCAC 89
QY 2062 AGCTGATACCAATCAAAACGAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAACTGA 2121
Db 90 AAATGATCAAGAAAAACCAACCAACCAAGAGAACCCAGCCCTTAAGAAAAACCTCT 149
QY 2122 GGAAGAAACCCCTCGAGAGAGAGAAACGGCAACCGGAAACCGAGAAACCGAGTCTCCA 2181
Db 150 TGATGACATAACTGTTGACCTGCACTGATGTAGTCTCTGAACTCTAGAGCCAA 209
QY 2182 AGAGGAACCAAGAAATCAACGAGGAATCAGAGGAATCAGAGAAACCTCAGGTGAGACTG 2241
Db 210 ACCAGAACCAAGAGCCGAAACCAAGAACCAAGAACCGAACCGAACCGAACCGAGGT 269
QY 2242 TGAAGAAAAACTG 2254
Db 270 TGAACCAAAAGTG 282

```

```

RESULT 21
US-08-956-171E-59
; Sequence 59, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
;
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986

```

Db	492	TAAGAGCAAGTGGCTACGGATACACAGGCTGTGTGATGAACACACAAAAAGTAGTGGATCA	551
Qy	2008	AGATAATGGTTTGGTAAGCGCTAGCGACCATGTTCAAAGAAACAAATAATGCTCAAGCTGA	2067
Db	552	AGCTCAACAGATGTTAACCAACACAGCTGTGTGTGATGAAAAAGCAAAAGAAACGAA	611
Qy	2068	TACCAATCAAAACGGAAAAACCAAGCGAGGAGAAACCTCGACACAGAAAAACCTGAGGAGA	2127
Db	612	TGCTGCTAAAGTGCACAAATGAAAAAGATCAACAAGCAGTAAACAGCTCGGAAACAGACACA	671
Qy	2128	AACCCCTCGAAGAGAAACCGCAAGCGAGAAACACAGAGTCTCCAAAAACCAACAGAGGA	2187
Db	672	AGCCAAGCTTGAAGAATTAGCGAAAAATCGGAAGCGGAAAAAGCAAAAGGCAGAAAAAGA	731
Qy	2188	ACCAGAAGAATCACAGAGGAATCAGAAACCTCAGGTCGAGACTGAAAAGGTTGAGA	2247
Db	732	ACAAGCAGCAAAAGAGCTGAATCGCTTAACAAACAGAAAGAGAAGCCAAAGCTAAAGA	791
Qy	2248	AAAACTCGAGAGGCTGAAGATTTACTTGGAAA	2280
Db	792	TCAAAAGCAAGGACGATCAGCAGTCGAGA	824

RESULT 23
 US-09-627-122-21/c
 ; Sequence 21, Application US/09627122
 ; Patent No. 6472521
 ; GENERAL INFORMATION:
 ; APPLICANT: Uhlmann, Eugen
 ; APPLICANT: Greiner, Beate
 ; APPLICANT: Unger, Eberhard
 ; APPLICANT: Gothe, Gislinde
 ; APPLICANT: Schwerdel, Marc
 ; TITLE OF INVENTION: OLIGONUCLEOTIDES FOR THE INHIBITION OF HUMAN e95
 ; TITLE OF INVENTION: EXPRESSION
 ; FILE REFERENCE: 02481.1678
 ; CURRENT APPLICATION NUMBER: US/09/627,122
 ; CURRENT FILING DATE: 2000-07-27
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 5340
 ; TYPE: DNA
 ; ORGANISM: Plasmodium falciparum
 US-09-627-122-21

Query Match 2.0%; Score 45.4; DB 4; Length 5340;
 Best Local Similarity 44.3%; Pred. No. 0.0056;
 Matches 178; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

Qy	9	ATCAACTTGGTTCGTCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTTCTTTATA	68
Db	928	ATAATAATTATCTTTATTAATGATGATAATAAAATGGATCAAGGTCATGATCATATAAAT	869
Qy	69	TAGATGGTGTATCAGGCTGGTCAAAAGGCAGAAAACTTGACACACAGATGAAGTCAGTAAGA	128
Db	868	ATCTTTATATGATGATAATAAATGGATCAAGGTCATGATCATATAATATCTTTATA	809
Qy	129	GGAGGGGATCAACGCCGAAACAAATNGTATCAAGATTACGGATCAAGGTTATGTGACCT	188
Db	808	ATGATGATAATAAAATGGATCAAGGTCATCATATAATTAATTTATCTTTATAATGATGATA	749
Qy	189	CTCATGGAGACCAATTATCATCTATATATGGCAAGGTTCCCTTATGATGCCATCATCAGTG	248
Db	748	ATAAATAGATCAAGGTGATCATCAATAATTATCTTTATAATGATGATATCATTAAGAC	689
Qy	249	AGAGAGCTCTCATGAAGAATCCGAATTTATCAGTTGAAGGATTCAGACATTTGCAATGAAA	308
Db	688	AAAAATTTACTTGGAGTTAAAAATACCTTGATGAAGAGATAGAACAAATTACACAA	629
Qy	309	TCAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGTNTACCTTAAAGGATCGAG	368
Db	628	CACATGATATAGATTTAAGTAAAGTATTTTGAACAAATCGATGAAGAACTTCGAAATTTCTTA	569

[illegible]

RESULT 24
US-08-072-610-1
; Sequence 1, Application US/08072610
; Patent No. 5532133
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens, and Diagnostic Assays
; TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,610
; FILING DATE: 19930602
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 5986/07686
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)527-7700
; TELEFAX: (212)753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
; IMMEDIATE SOURCE:
; CLONE: PYMB3.3.1

RESULT 25
US-08-719-822B-1
; Sequence 1, Application US/08719822B
; Patent No. 5874527
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/719,822B
; FILING DATE: 09/30/96
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 5986/17686US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)527-7700
; TELEFAX: (212)753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
; IMMEDIATE SOURCE:
; CLONE: pVMB3.3.1
US-08-719-822B-1

RESULT 26
US-09-092-458-1
; Sequence 1, Application US/09092458

```

; Patent No. 6231861
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
; TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,458
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/719,821
; FILING DATE: 09/30/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Cogoris, Adda
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 5986/17686US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)527-7700
; TELEFAX: (212)753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
; IMMEDIATE SOURCE:
; CLONE: pVMB3.3.1
; US-09-092-458-1

Query Match          2.0%; Score 44.8; DB 3; Length 3337;
Best Local Similarity 50.0%; Pred.No.0.0062;
Matches 112; Conservative 0; Mismatches 112; Indels 0; Gaps 0

QY      2066  GATACCAATCAAACGGAAACAACGCGAGGAGAAAACCTTCAGACAGAAAAAACCTGGA
Ddb      |||
QY      2070  GAAACTGCAGAAGCAGAAGAAAGTGGGAAGAGGTACCTGCAGAGTAGAAGAAAGTGCA
Ddb      |||
QY      2126  GAAACCCTCGAAGAGAAACCGCAANGCGAGAACACAGAGTCTCCAAACACCAAC
Ddb      |||
QY      2130  GTACCTCGAAGTAGAAGAGTGGGAAGAGGTACCAAGAAAGTAGAAGAGGTACCA
Ddb      |||
QY      2186  GAACCAAGAANAATCACACAGAGGAATCAGAAAGAACCTTCAGGTTCGAGACTGAAAAGGT
Ddb      |||
QY      2190  GAAGTAGAAGAAAGTGGGAAGAGGTACCAAGAAAGTGGGAAGAGTTAGAAAGGTACCA
Ddb      |||
QY      2246  GAAAAACTGACAGAGGCTGAAGATTTCATTGGAATAATCCAGGA   2289
Ddb      |||
QY      2250  GAGGTACCAAGAAAGTGGGAAGAGGTACCAAGAAAGTGGGAAGTGGGAAGTGGGAAGT
Ddb      |||

```

RESULT 27
US-08-742-185-101
; Sequence 101, Application US/08742185
; Patent No. 6020476

```

/ GENERAL INFORMATION:
/ APPLICANT: Page, David C.
/ APPLICANT: Reijo, Renee
/ APPLICANT: Saxena, Richa
/ APPLICANT: Hawkins, Trevor
/ APPLICANT: Reeve, Mary Pat
/ TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
/ NUMBER OF SEQUENCES: 102
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
/ STREET: Two Militia Drive
/ CITY: Lexington
/ STATE: Massachusetts
/ COUNTRY: US
/ ZIP: 02173
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/742,185
/ FILING DATE: 30-OCT-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/690,734
/ FILING DATE: 31-JUL-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/310,429
/ FILING DATE: 22-SEP-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Granahan, Patricia
/ REGISTRATION NUMBER: 32,227
/ REFERENCE/DOCKET NUMBER: WHI94-07A2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 861-6240
/ TELEFAX: (617) 861-9540
/ INFORMATION FOR SEQ ID NO: 101:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 43795 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-742-185-101

```

RESULT 28
US-08-914-479A-3
; Sequence 3, Application US/08914479A
; Patent No. 6419332
; GENERAL INFORMATION:
; APPLICANT: Dale, James R.

;; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
;; FILE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
;; FILE REFERENCE: 481112.404C2
;; CURRENT APPLICATION NUMBER: US/08/914,479A
;; CURRENT FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: 08/409,270
;; PRIOR FILING DATE: 1995-03-23
;; PRIOR APPLICATION NUMBER: 07/945,860
;; PRIOR FILING DATE: 1992-09-16
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 765
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: An antigen of M5 and a carrier of the
;; OTHER INFORMATION: COOH-terminal portion of M5
US-08-914-479A-3

Query Match 1.9%; Score 43.8; DB 4; Length 765;
Best Local Similarity 48.2%; Pred. No. 0.0045;
Matches 123; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 1961 GCGACTGTCAAGTACTATGTGCAACATCCAAACGATCCGCAATTCAGATAATGTTTT 2020
DB 280 GCAAGTCGAAAGGCTTCGCGTGAATTAGACGATCAGTGAAGCTAAGAAACAAAGTT 339
QY 2021 GGTAAAGCTAGCGACCATGTTCAAGAAACAAATAATGTTCAAGCTGATACCAATCAAACG 2080
DB 340 GAAAGAGCTTTAGAGAGCAAGCAACAGCAAAATAGTGTCTTTGAAAGAACTTAACAAAGAG 399
QY 2081 GAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAAACCTTGAGAGAAACCCCTCGAGAA 2140
DB 400 CTTGAGAGAAAGCAAGAAATTAACAGAAAGAAAGAAAGCTGACTACAGCAAACTTGA 459
QY 2141 GAGAAACCGAAAGCGAGAAACCGAGAGTCTCCAAACCAACAGAGAAACCGAGAAATCA 2200
DB 460 GCAGAGCAAAAGCACTCAAGAAACAATTAGCAAAACAGCTGAAGAACTTGCAAAACCTA 519
QY 2201 CCAGAGGAATCAGAA 2215
DB 520 AGAGCTGGAAAGCA 534

RESULT 29
US-08-914-479A-5
; Sequence 5, Application US/08914479A
; Patent No. 6419932
; GENERAL INFORMATION:
; APPLICANT: Dale, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; FILE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
; FILE REFERENCE: 481112.404C2
; CURRENT APPLICATION NUMBER: US/08/914,479A
; CURRENT FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 08/409,270
; PRIOR FILING DATE: 1995-03-23
; PRIOR APPLICATION NUMBER: 07/945,860
; PRIOR FILING DATE: 1992-09-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: An antigen of three segments of M5 and a carrier
; OTHER INFORMATION: of the COOH-terminal portion of M5
US-08-914-479A-5

Query Match 1.9%; Score 43.8; DB 4; Length 855;
Best Local Similarity 48.2%; Pred. No. 0.0049;

Matches 123; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 1961 GCGACTGTCAAGTACTATGTGCAACATCCAAACGATCCGCAATTCAGATAATGTTTT 2020
DB 370 GCAAGTCGAAAGGCTTCGCGTGAATTAGACGATCAGTGAAGCTAAGAAACAAAGTT 429
QY 2021 GGTAAAGCTAGCGACCATGTTCAAGAAACAAATAATGTTCAAGCTGATACCAATCAAACG 2080
DB 430 GAAAAAGCTTTAGAGAGCAAGCAACAGCAAAATAGTGTCTTTGAAAGAACTTAAACAAAGAG 489
QY 2081 GAAAAACCAAGCGAGGAGAAACCTCAGACAGAAAAACCTTGAGAGAAACCCCTCGAGAA 2140
DB 490 CTTGAGAGAAAGCAAGAAATTAACAGAAAGAAAGCTGACTACAGCAAACTTGA 549
QY 2141 GAGAAACCGAAAGCGAGAAACCGAGAGTCTCCAAACCAACAGAGAAACCGAGAAATCA 2200
DB 550 GCAGAGCAAAAGCACTCAAGAAACAATTAGCAAAACAGCTGAAGAACTTGCAAAACCTA 609
QY 2201 CCAGAGGAATCAGAA 2215
DB 610 AGAGCTGGAAAGCA 624

RESULT 30
US-08-937-271-9
; Sequence 9, Application US/08937271
; Patent No. 6063386
; GENERAL INFORMATION:
; APPLICANT: Dale, James B.
; APPLICANT: Lederer, James W.
; TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN
; TITLE OF INVENTION: VACCINE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,271
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman, Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 481112.405C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 918 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..915
US-08-937-271-9

Query Match 1.9%; Score 43.8; DB 3; Length 918;

```

US-09-621-976-10407

Query Match      1.9%; Score 43; DB 4; Length 478;
Best Local Similarity 51.9%; Pred. No. 0.0058;
Matches 122; Conservative 0; Mismatches 110; Indels 3; Gaps 1;

QY      2046  GAAACAAAAATGCTCAAGCTGATACCAATCAAACGGAAAAACCAACGCGAGGAGAAACCTC 2105
Db      7    GAAAGGAAATGGAGAGAGATGGAAAAAGAGAAAGGAGAAAGTGAATAAGAGGAGAGAGACA 66

QY      2106  AGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGACGAGAAACCCAG 2165
Db      67  GAAAAGAAAACAGGAGATGGAAAAGAGAAATCAAGATGGAAAAGAGAAAGCGGAGATAAAAAAG 126

QY      2166  AGTCTCCAAAACCAACAGAGGAAACCGAGAAATCACCGAGGGAATCAGAGAGAACCTCAGG 2225
Db      127  AGGGGAAGAGATGTAAAGTCAAGAAGATGAAAA---AGAGAGAGAGATGGGAAAGAAG 183

QY      2226  TCGAGAGACTGAAAAGGTTGAAGAAAAAATCGAGAGAGGCTGAAGAGATTTACTTTGGAAA 2280
Db      184  ATCAAGGTGGAATGAGGAAGAAGCTGAAAAGAGAGAAAGAAAGATTTAAAAAGAGA 238

```

RESULT 33
US-09-286-981B-24
; Sequence 24, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743

```

? NUMBER OF SEQ ID NOS: 38
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 24
? LENGTH: 1284
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence:cdna derived
? OTHER INFORMATION: from the genome of Streptococcus pneumoniae
? US-09-286-981B-24

```

DB	913	GAACCTTGAAATATGCTGAGAGCTCGGAGAGTTATATAAAATACCGGAGCTTGGACCTAGTAA	972
QY	2069	ACCAATCAAACGGGAAAAACCAAGCGAGGAGAAAACCTCAGACAGAAAAACCTCGAGGAGAA	2128
DB	973	GAGGAAGCTAAGGAACTCTGAAACGAGGAAAAAGTTAAGCAAGCAAAAGCGGAGTTGAG	1032
QY	2129	ACCCCTCGAGAGAGAAACCGCAAGCGAGAAACCAAGAGTCTCCAAACCAACAGAGGAA	2188
DB	1033	AGTAAACAAGCTGAGGCTACAAGGTTAGAAAAATCAAGAGACAGATCGTAAAAAAGCAGAA	1092
QY	2189	CCGAGAAATCACGAGAGGAAATCAGAGAAACCTCAGGTCGAGACTGAAAAGGTTGAGAA	2248
DB	1093	GAAGAAGCTAAAACGAAAAGCAGCAGAAGAGATAAAGTTAAGAGAAAAACCAAGCTGAACA	1152
QY	2249	AAACTGAGAGAGGCTGAAGATTTACTTGGAAAA	2281
DB	1153	CCACAACGAGCGGGCTCTCTCAACAGAGAAAA	1185

RESULT 34

US-09-671-317-14/c
; Sequence 14, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62 US3 CIP
; CURRENT APPLICATION NUMBER: US/09/671,317
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm
; SEQ ID NO 14
; LENGTH: 929
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-454-363 : polymorphic base A or G
; NAME/KEY: misc.binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-454-363.mis1, potential
; NAME/KEY: misc.binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-454-363.mis2, potential complement
; NAME/KEY: primer bind
; LOCATION: 139..158
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer bind
; LOCATION: 634..652
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc.binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-454-363 potential probe
; NAME/KEY: misc.feature
; LOCATION: 674..679,881..882,892..893
; OTHER INFORMATION: n=a, g, c or t
US-09-671-317-14

Query Match 1.9%; Score 42.4; DB 4; Length 929;
Best Local Similarity 52.4%; Pred. No. 0.014;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 2115 AACCTGAGGAAGAACCCCTCGAGAGAGAAACCGCAAGCGAGAACCCAGAGTCTCCAA 2174
DB 896 AGCNVCACGGGANNCCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
QY 2175 AACCAACGAGGAACCCAGAGGAATCACCAGAGGAATCAGAGAACCTCAGGTGCGAGACTG 2234
DB 836 AAG 777
QY 2235 AAAGTTGAGCAAAAACCTGAGAGAGGCTGAAGATTACTTGAAAAA 2282
DB 776 AAG 729

RESULT 35

US-09-286-981B-34
; Sequence 34, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:

; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cdna derived
; OTHER INFORMATION: from the genome Streptococcus pneumoniae
US-09-286-981B-34

Query Match 1.9%; Score 42.4; DB 4; Length 1236;
Best Local Similarity 54.5%; Pred. No. 0.017;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2083 AAAACCAAGCGAGGAGAAACCTCAGACAGAGAAACCTTGAGGAGAAACCCCTCGAGAAGA 2142
DB 1080 AAAAGCAGCAGAGAGAGATTAAGTTAAAGAAACCAAGCTGAACACCAACCAAGCGCC 1139
QY 2143 GAAACCGCAAGCGAGAGAAACCGAGAGTCTCCAAAACCAACGAGGAAACGAGAGATCACC 2202
DB 1140 GGCTCTCTCAACCAAGAAACCAAGCTCCAGCTCCAAAACCAAGAGAAATCCAGCTGAACACC 1199
QY 2203 AGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAA 2238
DB 1200 AAAAGCAGAAAAACCAAGCTGATCAACAAAGCTGAAGA 1235

RESULT 36

US-09-286-981B-23
; Sequence 23, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cdna derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-23

Query Match 1.9%; Score 42.4; DB 4; Length 1338;
Best Local Similarity 54.5%; Pred. No. 0.018;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2083 AAAACCAAGCGAGGAGAAACCTCAGACAGAGAAACCTTGAGGAGAAACCCCTCGAGAAGA 2142
DB 1182 AAAAGCAGCAGAGAGAGATTAAGTTAAAGAAACCAAGCTGAACACCAACCAAGCGCC 1241
QY 2143 GAAACCGCAAGCGAGAGAAACCGAGAGTCTCCAAAACCAACGAGGAAACGAGAGATCACC 2202
DB 1242 GGCTCCAAAACCAAGAAACCAAGCTCCAGCTCCAAAACCAAGAGAAATCCAGCTGAACACC 1301

QY 2203 AGAGGAATCAGAAACCTCAGGTCGAGACTGAAAA 2238
Db 1302 AAAAGCAGAAAAACCCAGCTGATCACAAGCTGAAGA 1337

RESULT 37

US-09-221-017B-27
; Sequence 27, Application US/09221017B

; Patent No. 6444799

; GENERAL INFORMATION:

; APPLICANT: Ross, Bruce C.

; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

; NUMBER OF SEQUENCES: 1120

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 755 PAGE MILL ROAD

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: IBM Compatible

; SOFTWARE: FastSeq for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/221,017B

; FILING DATE: 23-DEC-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P1182

; FILING DATE: 31-DEC-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P1546

; FILING DATE: 30-JAN-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P2911

; FILING DATE: 09-APR-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/AU98/01023

; FILING DATE: 10-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Monroy, Gladys H

; REGISTRATION NUMBER: 32,430

; REFERENCE/DOCKET NUMBER: 27340-20021.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-813-5600

; TELEFAX: 650-494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5295 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: UNKNOWN

; ORIGINAL SOURCE:

; ORGANISM: PORPHYROMONAS GINGIVALIS

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1...5295

US-09-221-017B-27

Query Match 1.9%; Score 42.4; DB 4; Length 5295;

Best Local Similarity 54.5%; Pred. No. 0.047;

Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2093 GAGGAGAAACCTCAGACAGAAAAACCTGAGAGAAACCCCTCGAGAGAGAAACCGCAA 2152

Db 333 GAGATAGCCCAAGGAAGAAAAACCAACGACAGAGAGCTTAAACGGGAGGAAACCCAAA 392

QY 2153 AGCGAGAAACACAGAGTCTCCAAAAACCAACAGAGAACCAAGAAATCACCAGAGGAATCA 2212
Db 393 CCGGAGGAGCCCTGTACAGCGGCTCCGCTTGCCAGCCGGTCGAAAAACCTGTGGTAAA 452

QY 2213 GAAGAACTCAGGTCGAGACTGAAAAGGTTGAAGAA 2248

Db 453 CCACAGCAACCGGTCTATGACACAGAAACCTCAAGAA 488

RESULT 38

US-09-543-681A-1562

; Sequence 1562, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 1562

; LENGTH: 2766

; TYPE: DNA

; ORGANISM: Proteus mirabilis

US-09-543-681A-1562

Query Match

Best Local Similarity 1.8%; Score 42.2; DB 4; Length 2766;

Matches 122; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 2035 CCATGTTCAAGAAACAAAAATGGTCAAGCTCATACCAATCAAAACGGAACCAAGCGA 2094

Db 420 CGAAGCTGAAGAAAGCAAAACGTCGAAGCTGAAGAGCCCAAGAAAGGCTGAAGAAA 479

QY 2095 GGAGAAACCTCAGACAGAAAAACCTGAGGAGAAACCCCTCGAGAGAGAAACCGCAAAG 2154

Db 480 GGCAGAAACGTTGAAGCTGAAGAGCAAAACGTCGAAGCAGCGGAATTAGCTAAGCGCGAAGC 539

QY 2155 CGAGAAACACAGAGTCTCCAAACCAACAGAGCAAGCAAGCAATCACCAGAGCAATCAGA 2214

Db 540 AGCGGAAAAAGACAAAGTGAAACAAACAGAGAAACCAAAAGCTGATAAGCTGATCAGGA 599

QY 2215 AGAACCTCAGGTCGAGACTGAAAAAGGTTGAGAAAAAATCTGAGAGAGGCTGAAGATTACT 2274

Db 600 AAAGCACGTCGCATAGCGCAAGAGCTGAACCTGAAACGTTAAACAGAGAAAGCTCAGCG 659

QY 2275 TGGAAAAATCCAGGA 2289

Db 660 TCGTAAAGCTGAAGA 674

RESULT 39

US-09-286-981B-22

; Sequence 22, Application US/09286981B

; Patent No. 6503511

; GENERAL INFORMATION:

; APPLICANT: Wizemann, Theresa M.

; APPLICANT: Koenig, Scott

; APPLICANT: Johnson, Leslie S

; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines

; FILE REFERENCE: 469201-396

; CURRENT APPLICATION NUMBER: US/09/286,981B

; CURRENT FILING DATE: 1999-04-06

; PRIOR APPLICATION NUMBER: US 60/085,743

; PRIOR FILING DATE: 1998-05-15

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 22

; LENGTH: 1239

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cdna derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae.
US-09-286-981B-22

Query Match      1.8%; Score 42; DB 4; Length 1239;
Best Local Similarity 50.0%; Pred. No. 0.023;
Matches 105; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2039 GTTCAAGAAACAAATAAGTGTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAG 2098
Db 1030 GATCTTAAAGAGCAGAGAGAGAGCTTAACGAAAGCAGCAGAGAGAGATTAAGTTAA 1089
QY 2099 AAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGAG 2158
Db 1090 GAAAAACAGCTGACACACCAACACAGCGCGGCTCTCAACCGAGAAAAACCACTGAA 1149
QY 2159 AAACAGAGTCTCCAAACCAACAGAGGAAACAGAGAAATCAACAGAGGAATCAGAGAA 2218
Db 1150 GAGCTGAGATCCAGCTCCAGCTCCAAACCAAGAGAGAGAGCTGAACAAACCAAAAGCA 1209
QY 2219 CCTCAGGTCGAGACTGAAAGGTTGAAGAA 2248
Db 1210 GAAAAACAGATGATCAACAGCTGAAGAA 1239

RESULT 40
US-09-286-981B-28
; Sequence 28, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cdna derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-28

Query Match      1.8%; Score 42; DB 4; Length 1275;
Best Local Similarity 50.0%; Pred. No. 0.023;
Matches 105; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2039 GTTCAAGAAACAAATAAGTGTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAG 2098
Db 1066 GATCTTAAAGAGCAGAGAGAGAGCTTAACGAAAGCAGCAGAGAGAGATTAAGTTAA 1125
QY 2099 AAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGAG 2158
Db 1126 GAAAAACAGCTGACACACCAACACAGCGCGGCTCTCAACCGAGAAAAACCACTGAA 1185
QY 2159 AAACAGAGTCTCCAAACCAACAGAGGAAACAGAGAAATCAACAGAGGAATCAGAGAA 2218
Db 1186 GAGCTGAGATCCAGCTCCGCGACCAAAACCAAGAGAGAGAGCTGAACAAACCAAAAGCA 1245
QY 2219 CCTCAGGTCGAGACTGAAAGGTTGAAGAA 2248
Db 1246 GAAAAACAGATGATCAACAGCTGAAGAA 1275

RESULT 41
US-09-461-697-193
; Sequence 193, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-193

Query Match      1.8%; Score 41.4; DB 3; Length 696;
Best Local Similarity 51.5%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 2046 GAAACAAAAATGTCAGGCTGATACCAATCAACGCGAAACCAAGCGAGGAGAAACCTC 2105
Db 392 GAAAGGAAATGAGAGAGATGGAAGAGAGAGAGAGAGATGAAAGAGCGAGAGAGACA 451
QY 2106 AGACAGAAAAACCTGAGGAGAGAAACCCCTCGAGAGAGAGAAACCCCAAGCGAGAAACCG 2165
Db 452 GAAAGAGAAACAGGAGTTGGAAAAAGAGAGATGAAGATGGAAGAGAGAGAGAGATGAAAAAG 511
QY 2166 AGTCTCCAAAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225
Db 512 AGGGGAAAGATGTAAAGTCAAGAGAGAGATGAAA---AGAGAGAGAGAGATGGAAGAGAG 568
QY 2226 TCAGAGACTGAAAGAGTTGAAGAAAAAATCGAGAGAGGCTGAAGATTTTACTTTGGAAA 2280
Db 569 ATGAGGTGGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTTAAAGAGAG 623

RESULT 42
US-09-461-697-191
; Sequence 191, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-191
```

Query Match	1.8%;	Score 41.4;	DB 3;	Length 699;
Best Local Similarity	51.5%;	Prod. No. 0.023;		
Matches 121;	Conservative	0;	Mismatches 111;	Indels 3; Gaps 1;
2046	GAACAAAAAATGCTCAAGCTGATACCAATCAACCGAAAAAACCAAGCGAGGAGAAACCTC	2105		
395	GAACAGGAATGAGAGATGTGAAGAAGAAACGGAACATGAAAAGGAAGAAAGACA	454		
2106	AGACGAAAAAATCTGGAGAGAAACCCCTCGAGAAGAGAAACCGCAAAAGCGAGAAACCG	2165		
455	GAAGAAGAAACAGGAGTTTGGAAAACAGAGAAATGAAGATGGAAAAAGAGAAGCGAGATAA	514		
2166	AGTCTCCAAAACCAACAGAGGAACCGAAGAAGATCACCAAGAGAAATCAGAAGAACCTCAGG	2225		
515	AGGGGAAGAGTGTAAAGTCAAGAGAAGATGAAAA---AGAGAGAGCAAGATGGAAGAAGAG	571		
2226	TCGAGACTGAAAAGGTTGAAAGAAAAATCGAGAGGGCTGAAGATTTCCTTGGAAA	2280		
572	ATCGAAGGTGGAATGAGGAAGAGAGCTGGAAAGAGAAAGAAAGAGATTTAAAGAAAGA	626		

```

RESULT 43
US-09-461-697-189
; Sequence 189, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: LG, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-189

```

	Query Match	1.8%	Score 41.4	DB 3	Length 717
	Best Local Similarity	51.5%	Pred. No. 0.024		
	Matches 121	Conservative 0	Mismatches 111	Indels 3	Gaps 1
QY	2046	GAACAAAAATGGTCTCAAGCTGATACCAATCAAAACGGAAAAACCAACGAGGAGAAACCTC	2105		
DB	413	GAAAAGGAATCGAGAGATGGAAGAAGAGAAAGCGAAGNTGAAAAGAGGAAGACAGACA	472		
QY	2106	AGACAGAAAAACCTGAGGAAGAAAACCCCTCGAAGAGAGAAACCGCAAAACGCGAAGAACCCAG	2165		
DB	473	GAANAAGAAAACGAGAGTTGGAAAAGAGAATGAAGATGGAAAAGAGAAAGCGGAGATATAAAAAG	532		
QY	2166	AGTCTCTCAAAACCAACAGAGGAACACAGAGAAATCACCAGAGAGAAATCAGAAAGAACCTCAGG	2225		
DB	533	AGGGGAAGAGTGTAAAGTCTAAGAGAGATGAAAA--AGAGAGAGAGATGGAAGAGAGAG	589		
QY	2226	TCGAGACTGAAAAGGTTGAAGAAAAACTGAGAGAGCGCTGAAGATTTACTTGGAAA	2280		
DB	590	ATGAAGGTGGAATGAGGAGAGAGCTGGAAAAGAGAAAGAAAGAGATTTTAAAGAAGA	644		

RESULT 44
US-09-461-697-187
; Sequence 187, Application US/09461697
; Patent No. 6277974

```

; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Purbanum, Stuart D.
; APPLICANT: Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-461-597-187

```

Query Match.	1.8%;	Score 41.4;	DB 3;	Length 774;
Best Local Similarity	51.5%;	Pred. No. 0.025;		
Matches 121;	Conservative	0;	Mismatches 111;	Indels 3; Gaps 1;
QY	2046	GAACAACAAAATGCTCAAGCTGATACCAATCAACACGGAAACCAACGAGGAGAAACCTC	2105	
DB	470	GAACGGAAATGAGAAAGATGGAAAGAGAGAAAGGAGAGATGAAAAGAGGAAGAAGACA	529	
QY	2106	AGACAGAAAACCTGAGGAAGAAAACCCCTCGAGAGAGAGAAACCGCAAGCGAGAAACCCAG	2165	
DB	530	GAAGAAGAAACAGGAGTTGGAAAAGAGAAATCAAGATCGAAAAGAGAAAGCGGAGATATAAAAAG	589	
QY	2166	AGTCTCCAAAAACCAACAGAGGACACGAGAGATCAACGAGGGAATCAGAGAGACCTCAGG	2225	
DB	590	AGGGGAAGAGNTGTAAAAGTCAAGAAGNTGAAA---AGAGAGAGAAAGTGGAAAAGAAAG	646	
QY	2226	TCGAGACTGAAAAGGTTGAAGAAAAACGAGAGAGGCTGAAGATTTTACTTTGGAAA	2280	
DB	647	ATCAAGGTGGAATGAGGAGAAAGCTCGAAAAGAGAAAGAAAGATTTTAAAGAAGA	701	

RESULT 45
US-09-461-697-185
; Sequence 185, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-185

Query Match 1.8%; Score 41.4; DB 3; Length 819;
Best Local Similarity 51.5%; Pred. No. 0.026;
Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

		2046	GAAACAAAAATGCTCAAGCTGATACCATTCAAACGGAAAAAACCACGGAGGAGAACCCTC	2105
Qy				
		515	GAAAAGGAATTCGAGAAAGATGTGAAAAGAGAAAGAGACAAGATGATAAAAGAGGAAGAAGACA	574
Db				
		2106	AGACAGAAAAAACCTGAGGAGAGAAACCCCTCGAAGAGAGAAACCGCAACCGCGAAACCAG	2165
Qy				
		575	GAAAAGGAACAGAGATTTGGAAAAGAGANTGAAGATGGAAAAGAGAGAGGGAGATTAATAAAG	634
Db				
		2166	AGTCTCCAAAACCAACAGAGAGGAACCGAGAAGATCACACAGAGGAATCAGAGAAGAACCTTCAGG	2225
Qy				
		635	AGGGGAAGATGTAAAGTCAAAGAAGATGAAAA--AGAGAGAGAGATGGAAAAGAG	691
Db				
		2226	TCGAGACTGAAAAGCTTTGAAGAAAAAACTGAGAGAGCGCTGAAGATTTACTTCGGAAA	2280
Qy				
		692	ATCAAGGTGGAAATGAGGAAGAGCTGGAAAAAGAGAAAGAAAGATTTTAAAAGAAGA	746
Db				

RESULT 46
US-09-671-317-439/c
; Sequence 439, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62.US3.CIP

	Query Match	1.8%;	Score 41.4;	DB 4;	Length 1001;
	Best Local Similarity	49.8%;	Pred. No. 0.03;		
	Matches 123;	Conservative 0;	Mismatches 122;	Indels 2;	Gaps 1;
QY	2035	CCATGTTCAAGAAACAAAAATGGTCAGCTGATCCATCAAAACCGAAAAACCAAGCGA	2094		
Db	1001	CCATCTCAAGAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	942		
QY	2095	GGAGAACCTCAGACAGAAAAAACCCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAG	2154		
Db	941	AGAAAGAAGAGAAGAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	882		

2155	CGAGAAACGAGAGTCTCCAAAACCAACACAGAGGACACAGAAAGAAATCACACGAGAGAAATCAGA	2214
Qy		
881	AGAAGAAGAAGAAGAAGAAGAAGGA--ACAAGAAGAAGAAGAAGAAGAAGAAGAAGA	824
Db		
2215	AGAACCTCAGGTCGAGACTTGAAAAGGTTCAAGAAAACCTGAGAGAGGCTTGAAAGATTACT	2274
Qy		
823	AGAAGAAGAAGAAGAAGAAGAAGANNNNNNNAGAAGAAGAAGATGSCCGTGTGACTCGAAG	764
Db		
2275	TGGAATA	2281
Qy		
763	TGACACA	757
Db		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match	1.8%;	Score 41.4;	DB 3;	Length 1360;
Best Local Similarity	51.3%;	Pred. No. 0.037;		
Matches	96;	Conservative 0;	Mismatches 91;	Indels 0; Gaps 0;
QY	2083	AAAAACCAAGCGAGGAGAACCTCAGACAGAAAACCTGAGGAGAGAAACCCCTCGAGAAGA	2142	
Db	574	AAAGCGCAGAGAGAGATAAAGTTAAAGAAAAACAGCTGAACCAACCAACCGCGCC	633	
QY	2143	GAACCGCAAGCGAGAAACCAAGAGTCTCCAAACCAACAGAGGAAACAGAGAATCACC	2202	
Db	634	GGCTCCAAAGCAGAAAAACCAAGCTCCAGTCCAAAACCAAGAGAATCCAGCTGAACAACC	693	
QY	2203	AGAGGAATCAGAGAAACCTCAGTGCAGACTGAAGAGGTTGAAGAAAAAATCTGAGAGAGGC	2262	
Db	694	AAAAGCAGAAAAAACCAAGCTGATCAACAGCTGAAGAAGACTATGCTCGTAGATCAAGAAGA	753	
OV	2263	TGAAGAT	2269	

Db 754 AGAATAT 760

RESULT 48

US-09-536-784-37

Sequence 37, Application US/09536784

Patent No. 6573082

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 1360 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 37:

US-09-536-784-37

Query Match 1.8%; Score 41.4; DB 4; Length 1360;

Best Local Similarity 51.3%; Pred. No. 0.037;

Matches 96; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

2083 AAACACGAGCGAGGAAACCTCAGACAGAAAACCTGAGGAGAAACCCCTCGAGAAGA 2142

574 AAACGACGAGAGAGAGATAAAGTTAAAGAAAACACAGCTGACACCAACCAACGCGCC 633

2143 GAAACCGGAAACGAGACACCAAGCTCCAAACCAACAGAGGACCAAGAAATCACC 2202

634 GGCTCCAAAGCAGAAAACCAAGCTCCAGCTCCAAACCAAGAGAAATCCAGCTGAACAAC 693

2203 AGAGGAATCAGAAAGACCTTCAGTCCAGTCCAGACTGAAAAGGTTGAAGAAAACCTGAGAGGCG 2262

694 AAACGAGAAAACCAAGCTGATCAACAAGTGAAGAGACTATGCTCGTAGATCAGAGA 753

2263 TGAAGAT 2269

754 AGAATAT 760

RESULT 49

US-09-461-697-184

Sequence 184, Application US/09461697

Patent No. 6277974

GENERAL INFORMATION:

APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Puranam, Kasturi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
TITLE OF INVENTION: CELL DEATH
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 184
LENGTH: 1669
TYPE: DNA
ORGANISM: Homo sapiens
US-09-461-697-184

Query Match 1.8%; Score 41.4; DB 3; Length 1669;

Best Local Similarity 51.5%; Pred. No. 0.043;

Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 2046 GAAACAAAATGGTCAAGCTGATACCAATCAACCGAAAAACCAAGCGAGGAGAAACCTC 2105

Db 594 GAAAGGAATGGAGAGATGAAAAGAGAAAGAGAGATGAAAAGAGGAGAGACA 653

QY 2106 AGACAGAAAAACCTGAGGAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAGAACCTCAG 2165

Db 654 GAAAAGAAACAGGAGTTGGAAAAGAGAAATGAAGTGGAAAAGAGAGAGGAGATATAAAAAG 713

QY 2166 AGTCTCAAACCAACAGAGAACCCAGAGAAATCACCAGAGGAATCAGAGAACTCAGG 2225

Db 714 AGGGGAAAGATGTAAGAACTCAAGAGAGATGAAAA---AGAGAGAGAGATGGAAGAGAG 770

QY 2226 TCGAGACTGAAAAGTTGAGAAAAACCTGACAGAGGCTGAAGATTTACTTGAAA 2280

Db 771 ATGAAGTGAATGAGGAGAGAGTGGAAAAGAGAAAGAGATTTTAAAAGAGA 825

RESULT 50

US-08-961-527-134

Sequence 134, Application US/08961527

Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 1, 2004, 06:21:35 ; Search time 776 Seconds
(without alignments)
4447.684 Million cell updates/sec

Title: US-09-765-272A-66

Perfect score: 4019

Sequence: 1 CSYELGRHOAGQVKKESNRV.....TEKVEEKLEAREDLGKIQD 763

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09765272/runat_30042004_081052_11633/app_query.fasta_1.903
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=150 -DOCLALIGN=200 -NORM SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US09765272 -CGN_1_232 @runat_30042004_081052_11633
-NCFU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99.9	4.013	2290	9	US-09-765-272-65
2	3968.5	98.7	2481	10	US-09-769-787-206
3	3919	97.5	2531	13	US-10-412-862-11
4	3919	97.5	2531	16	US-10-412-850-11
5	3919	97.5	2531	16	US-10-387-783-11
6	3665	91.2	2639	10	US-09-884-465A-5
7	3644.5	90.7	2531	13	US-10-412-862-5
8	3644.5	90.7	2531	16	US-10-412-850-5
9	3644.5	90.7	2531	16	US-10-387-783-5
10	3641.5	90.6	8195	13	US-10-158-844-94
11	3588.5	89.3	2523	10	US-09-884-465A-3
12	3588.5	89.3	2647	10	US-09-884-465A-4
13	2658.5	66.1	2451	13	US-10-412-862-9
14	2658.5	66.1	2451	16	US-10-412-850-9
15	2658.5	66.1	2451	16	US-10-387-783-9
16	2649.5	65.9	2389	9	US-09-765-272-55
17	2424	60.3	2359	13	US-10-158-844-243
18	1377	34.3	973	13	US-10-158-844-355
19	1228.5	30.6	3120	10	US-09-884-465A-1
20	1228.5	30.6	5048	10	US-09-884-465A-2
21	1202.5	29.9	6867	13	US-10-158-844-192
22	1200	29.9	1684	13	US-10-158-844-258
23	1167.5	29.0	1455	10	US-09-769-787-246
24	1167.5	29.0	1455	10	US-09-769-787A-23
25	1167.5	29.0	1455	13	US-10-412-862-7
26	1167.5	29.0	1455	16	US-10-412-850-7
27	1167.5	29.0	1455	16	US-10-387-783-7
28	1141.5	28.4	1342	9	US-09-765-272-181
29	1023.5	25.5	5215	10	US-09-252-088-13
30	1023.5	25.5	5215	16	US-10-340-792-13
31	1019.5	25.4	2469	10	US-09-769-736-17
32	980	24.4	2528	10	US-09-884-465A-9
33	877	21.8	1146	10	US-09-769-736-23
34	827.5	20.6	819	10	US-09-884-465A-257
35	827.5	20.6	819	16	US-10-324-143-10
36	827.5	20.6	2721	16	US-10-324-143-6
37	822.5	20.5	816	16	US-10-324-143-4
38	822.5	20.5	816	16	US-10-324-143-5
39	542.5	13.5	1455	10	US-09-769-736-71
40	512	12.7	841	9	US-09-452-599-34
41	512	12.7	841	15	US-10-121-120-34
42	190.5	4.7	3466	16	US-10-104-047-536
43	186.5	4.6	3494	13	US-10-087-192-233
44	186.5	4.6	5316	16	US-10-159-563-295
45	184.5	4.6	3611	9	US-09-974-298-113
46	180.5	4.5	5126	15	US-10-232-561-9
47	179.5	4.5	3570	16	US-10-369-493-45846
48	178	4.4	2522	9	US-09-842-553-98
49	177.5	4.4	8588	9	US-09-816-669A-13
50	177.5	4.4	8588	13	US-10-342-887-1335
51	177.5	4.4	8588	13	US-10-172-118-1335
52	175	4.4	3891	13	US-10-282-122A-35508
53	169.5	4.2	2232	15	US-10-087-464-45
54	169	4.2	3825	13	US-10-283-122A-7653
55	168.5	4.2	3284	15	US-10-128-714-7305
56	168.5	4.2	5142	13	US-10-087-192-230
57	168	4.2	2883	15	US-10-128-714-7193
58	168	4.2	2939	15	US-10-128-714-6193
59	168	4.2	4225	15	US-10-128-714-193
60	168	4.2	4939	15	US-10-128-714-5193
61	167.5	4.2	3336	9	US-09-815-242-8857
62	167	4.1	4088	10	US-09-814-353-20949
63	165	4.1	2392	9	US-09-216-393-331
64	165	4.1	2392	15	US-10-321-856-331
65	165	4.1	2393	9	US-09-216-393-329
66	165	4.1	2393	15	US-10-321-856-329
67	165	4.1	3379	15	US-10-128-714-6305
68	165	4.1	5379	15	US-10-128-714-5305
69	164.5	4.1	9072	16	US-10-396-122-87
70	164	4.1	2865	13	US-10-424-599-24135
71	164	4.1	3003	15	US-10-128-714-2305
72	164	4.1	3809	14	US-10-001-870-66

61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
182 GTGACCTCTCATGGAGACCATATCATATCATATAATGGCAAGGTTCCTTATGATGCCATC 241
81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
242 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAATATATCAGTTGAAGGATTGAGACATTTGTC 301
101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
302 AATGAATCAAGGGTGTATGTCTATTAAGTAAACGGTAAATATCTATGNTACCTTAAG 361
121 AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
362 GATGCAGCTCATGGGATAATATTCGGACAAAAGAGAGATTAAACGTCAGAGCAGGAA 421
141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
422 CGCAGTCAATATCAATACTCAAGAGCAGATAATGCTGTGTCGAGCCAGAGCCCAAGGA 481
161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
482 CGTTATACACGAGATGATGGGTATATCTTCATATGCATCTGATATCATTTGAGGACACGGGT 541
181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
542 GATGCTTATATGCTTCACGGCGACCATACCATATACATTCCTAAGAATGAGTTATCA 601
201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgProSer 220
602 GCTAGCAGTTAGTCTGCTCAGAGGCTATGGAATGGGAAGCAGGAGTCTCGTCTCTCT 661
221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
662 TCAAGTCTAGTTAATGCAATATCCAGTCAACCAAGATTGTACAGAGAACCAATCTG 721
241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
722 ACTGTCATCCCACTTATCATCAATCAAGGGAACCAATTCAGACCTTTTACGTGAA 781
261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
782 TTGTATGCTAAACCCCTTATCAGACGCCATGTGGAATCTGATGGCTTATTTTCGACCCA 841
281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
842 CGGCAAAATCACAAAGTCGAACCCGAGAGGTGTAGCTGTCCCTCATGTATACCAATTACCAC 901
301 PheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320
902 TTTATCCCTTATGAACAATGTCTGNAITGGAAAAACGAATGCTGTATATATCCCTTT 961
321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340
962 CGTTATCGTTCAAAACCATTTGGGTACAGATTCAGACCCAGAACCAACCAAGTCCCAATCG 1021
341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
1022 ACTCCGGAACTTAGTCCAAAGTCGCAACCTGCACCAAAATCCTCAACGAGCTCCAAAGCAAT 1081
361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
1082 CCANITGATGAGAAATTTGTCTAAAGAGCTGTTCGAAAAGTAGGCGCATGTTATGCTTT 1141
381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
1142 GAGGAGAAATGGATTCTCGTTATATCCAGCCCAAGATCTTTCACGAGAAACAGCAGCA 1201
401 GlyIleAspSerLysLeuAlaLysGlnGlnSerLeuSerHisLysLeuGlyAlaLysLys 420
1202 GGCATTGATAGCAAACTGGCCCAACGAGGAAAGTTTATCTCATAAAGCTAGGAGCTAAGAAA 1261

421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
1262 ACTGACCTCCCATCTAGTATGATCGAGAAATTTTACATAATAGGCTTATGACTTACTAGCAAGA 1321
441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
1322 ATTCAACCAAGATTTACTTTGATAATAAGGTCACAAAGTTGATTTTGGGCTTTTGGATAAC 1381
461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
1382 CTGTTGGACACACTCAAGATGTCTCAAGTATAAAGTCAAGTGTAGTGTAGGATATCTTT 1441
481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
1442 GCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAAAACCAATGCGCAATATACC 1501
501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
1502 TACACTGATGATGAGATTCAAGTAGCTCAAGTTGGCAGGCAAGTACACACAGAAAGCGGT 1561
521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
1562 TATATCTTTGATCTCTCGTATATACCAAGTATGAGGGGAGTCCATGTACTTACTCCACAT 1621
541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 560
1622 ATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGACGC 1681
561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
1682 CAGGCTTATGTCTAAAGAGAAAGGTTTGACCCCTCTCTGCACAGACCATCAGGATTCCAGGA 1741
581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysVal 600
1742 AATACTCAGGCAAAAGGAGCAGAGCTATCTACACCCGCTGAAAGCAGCTAAGAGGTG 1801
601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
1802 CCATCTGATCGTATGCTTACATCTTCAATATCTGTAGAGTCAAAAACCGTAGTTTA 1861
621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640
1862 ATCATACCTCATATGACCATTAACCAATCAATCAAAATTTGAGTGTGTTTACGAGGSCCTT 1921
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrVal 660
1922 TATGAGGCACCTAAGGGGTATACTCTTTGAGGATCTTTTGGCGACTGTCAAGTACTATGTC 1981
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
1982 GACATCCAAACGAAACGTCGCAATTCAGATTAATGTTTTGTTAGCTAGCGACCATGTT 2041
681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700
2042 CAAAGAAACAAAATGCTCAAGCTGATACCAATCAACCGAAAAACCAAGCGAGGAGAAA 2101
701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720
2102 CCTCAGACAGAAAAACCTGAGGAGAAACCCCTCGAGAAAGAGAAACCGCAAGCGAGAAA 2161
721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
2162 CCAGAGTCTCCAAAAACCAAGAGAAACCAAGAAATCACCAGAGGAATCAGAGAAACCT 2221
741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLys 760
2222 CAGGTCCAGACTGAAAAGGTTGAAGAAAACCTGAGAGGCGCTGAAGATTACTTGGAAAA 2281
761 IleGlnAsp 763
2282 ATCCAGGAT 2290

RESULT 2

```

US-09-769-787-206
; Sequence 206, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip W
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129W0
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 206
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-206

Alignment Scores:
Pred. No.: 0 Length: 2481
Score: 3968.50 Matches: 754
Percent Similarity: 99.08% Conservative: 3
Best Local Similarity: 98.69% Mismatches: 6
Query Match: 98.74% Indels: 1
DB: 10 Gaps: 1

US-09-765-272A-66 (1-763) x US-09-769-787-206 (1-2481)

QY 1 CysSerTyrGluLeuGlyValArgHisGlnAlaGlyGlnValLysGluSerArgVal 20
DB 58 TGTTCCTATAGCTGGTGGTCCACCAAGCTGGTCAGGATAAGAAAGAGTCTAATCGAGTT 117
QY 21 SerTyrIleAspGlyValAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
DB 118 GCTTATATAGATGGTGCATCAGCTGGTCAAAAGGCGAAGAACTTGACACCATGAGTTC 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
DB 178 AGTAGAGGGGGGATCAACCGCAACAAATCGTCTCATCAAGATTACGGATCAAGGTTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
DB 238 GTGACCTCTCATGGAGACCATTTATCTACTATATAATGGCAAGTCCCTTATGATGCCATC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysAspSerAspIleVal 100
DB 298 ATCAGTGAAGAGCTCTCATGAAGATCCGANTTATCAGTTGAAGGATTCAGACATTGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLys 120
DB 358 AATGAAATCAAGGGTGGTATGTCATCAAGCTAGGAGAGGAAATTAATCTGTTACCTTAAG 417
QY 121 AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
DB 418 GATGAGCTCATGCGGATTAATTCGACAAAGAGAGATTAACCGTCAAGAGCAGGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaArgAlaGln 159
DB 478 CGCAGTCATATCAACGGGTGAGAGCTAAGCATGATCATGCTAGTGTGACCCAGAGCCCAA 537
QY 160 GlyArgTyrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr 179
DB 538 GGACGCTATACAAAGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACAG 597
QY 180 GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu 199
DB 598 GGTGATGCTTATATCGTCTCCACGCGCCAGTACCATTTACATTTCTAAGAAATGAGTTA 657

200 SerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgPro 219
658 TCAGCTAGCGAGTAGCTGCTGCAGAAAGCCATTATGGAATGGAAAGCAGGAGTCTCGTCT 717
QY 220 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn 239
DB 718 TCTTCAAGTTCCTAGTTATAATCAATCCAGTCCAAAGATTGTCCAGAACCCCAAT 777
QY 240 LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259
DB 778 CTGACTGTCACTCACTCACTTATCATCAAAATCAAGGGGAAACATTTCAAGCTTTTACGT 837
QY 260 GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279
DB 838 GAATTTGATGCTAAACCCCTTATCAGAACGCCATCTGGAATCTGATGGCTTATTTTCGAC 897
QY 280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299
DB 898 CCAGCGCAATCAAGTCAAGCCGAGAGGTGTAGTCTCCCTCATGGTAAACCATTTAC 957
QY 300 HisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgGlnAlaArgIleIlePro 319
DB 958 CACTTTATCCCTTATGAAACAAATGTCTGAATGGAAAGAAAGAAATGCTCGTATTATTCCC 1017
QY 320 LeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGln 339
DB 1018 CTTTCGTTATCGTTCAAAACCATTTGGGTACCAAGATTCAAGACCAGAACCAACCAAGTCCCAA 1077
QY 340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359
DB 1078 TCAGCTCCGGAACCTAGTCCCAAGTCCGCAACCTCCACCAATCTCCACCAAGTCCCAAGC 1137
QY 360 AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal 379
DB 1138 AATCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTTCAAAAGAGAGGCGATGTTATGTC 1197
QY 380 PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla 399
DB 1198 TTTGAGGAGAAATGGAGTTTCTCGTTATATCCACCAAGGATCTTTCAGCAGAACAGCA 1257
QY 400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys 419
DB 1258 GCAGGCAATGATAGCAAACTGGCCCAAGCAGCAAGATTTATCTCATTAAGCTAGGAGCTAAG 1317
QY 420 LysThrAspLeuProSerSerAspAspGluPheTyrAsnLysAlaTyrAspLeuAla 439
DB 1318 AAACTGACCTCCCATCTAGTATCGAGAAATTTTACAAATTAAGGCTTATGACTTACTAGCA 1377
QY 440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp 459
DB 1378 AGAATTCACCAAGATTTACTTTGATAATAAGGTCGACAAAGTTGATTTTGAAGCTTTGGAT 1437
QY 460 AsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIle 479
DB 1438 AACCTGTTGGAAGCAGCTCAAGGATGTCCTCAAGTATGATAAGTCAAGTATGATGATGATATT 1497
QY 480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499
DB 1498 CTTGCTCTTCTAGTCCGATTCGATCCAGAAAGCTTTAGGAAACCAAAATTCGCGAAAT 1557
QY 500 ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp 519
DB 1558 ACCTACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1617
QY 520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro 539
DB 1618 GGTATATCTTTGATCCTCGTATATAACCAAGTATGATGATGATGATGATGATGATGATGAT 1677
QY 540 HisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 559
DB 1678 CATATGACCATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737
QY 560 AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer 579

```

```

1738 GCCCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCA 1797
580 GlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLys 599
1798 GGAATACTAGGCAAAAGAGAGAGAGATCTTACAAACCGCTGGAAGCAGCTAAGAAG 1857
600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619
1858 GTGCCACTTGATCGTATGCTTACAACTTCAATATCTAGAGGTCAAAACGGTAGT 1917
620 LeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGly 639
1918 TTAATCATCATCTCATATGACCATTAACCATTAACATAACATAAATTTGAGTGGTTTGACGAAGGC 1977
640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr 659
1978 CTTTATGAGGACCTAAGGGGTATCTCTTTGAGGATCTTTTGGCGACTGTCAAGTACTAT 2037
660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679
2038 GTGCAACATCCAAACGACGCTCGCATTCAGATAATGGTTTGGTAACGCTAGCGACCAT 2097
680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699
2098 GTTCAAAGAAACAAATAATGGTCAAGCTGATACCAATCAAAACGGAAAAACCAAGCGAGGAG 2157
700 LysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGlu 719
2158 AAACCTCAGACAGAAACAACTGAGGAAGAAACCCCTCGAAGAGAAACCGCAAGCGAG 2217
720 LysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGlu 739
2218 AAACGAGCTCTCCAAACCAACAGAGAACCGAGAGAAATCACAGAGGAAATCAGAGAA 2277
740 ProGlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGly 759
2278 CCTCAGGTCGAGACTGAAAGGTTGAAGAAAAAATCGAGAGAGGCTGAAGATTACTTTGGA 2337
760 LysIleGlnAsp 763
2338 AAAATCCAGGAT 2349

```

RESULT 3

```

JS-10-412-862-11
: Sequence 11, Application US/10412862
: Publication No. US20040052781A1
: GENERAL INFORMATION:
: APPLICANT: Johnson, Leslie S.
: APPLICANT: Adamou, John E.
: TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
: TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
: TITLE OF INVENTION: Motifs
: FILE REFERENCE: 469201-685
: CURRENT FILING DATE: 2003-04-14
: PRIOR APPLICATION NUMBER: 09/468,656
: PRIOR FILING DATE: 1999-12-21
: PRIOR APPLICATION NUMBER: 60/413,048
: PRIOR FILING DATE: 1998-12-21
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 3.0
: SEQ ID NO 11
: LENGTH: 2531
: TYPE: DNA
: ORGANISM: Streptococcus pneumoniae
JS-10-412-862-11

```

Alignment Scores:

```

Pred. No.: 0 Length: 2531
Score: 3919.00 Matches: 745
Percent Similarity: 98.03% Conservative: 3
Best Local Similarity: 97.64% Mismatches: 9

```

```

Query Match: 97.51% Indels: 6
DB: 13 Gaps: 1
US-09-765-272A-66 (1-763) x US-10-412-862-11 (1-2531)
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 58 TGTTCCTATGAGCTGGACGTTACCAAGCTGGTCAGGATAAGAAAGAGTCTAATCGAGTT 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 GCTTATATAGATGGTGAATCAGCGCTGGTCAAAAGGAGGAAACCTTGACACCATGAAGTC 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGTAAGAGGAGGGGATCAACGCCGACAAATTTGTATCAAGATTACGGATTCAAGGTTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTGACCTCTCATGGAGACCATTTATCATTTACTATATAATGGCAAGGTTCTTTATGATGCCATC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProSerTyrGlnLeuLysAspSerAspIleVal 100
Db 298 ATCAGTGAAGAGCTCCTCATGAAGATCCGATTATCAGTTCAAGGATTCAGACATTGTC 357
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 358 AATGAATCAAGGGTGGTTATGTCTAATGAAGTAAACGGTAATACTATTTTACCTTAAG 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGCGCTCATGCGGATAATATTCGGACAAAGAGAGATTAAACGTCAGAGACAGGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
Db 478 CGCAGTCATAATCAATACTCAAGAGCAGATAATGCTGTGTGTCGACGAGAGCCCAAGGA 537
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 538 CGTTATACACGAGATGAGGTATATCTTCATGATCTGATATCATTCAGGACACGGGT 597
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 598 GATGCTTATATCGTTCTTCACGGCGACCATTAACATTACATTCTTAAGAATGAGTTATCA 657
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220
Db 658 GCTAGCGAGTTAGCTGTGCGAAGCCATTGGAATGGGAAGCAGGGATCTCGTCTCTTCT 717
QY 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
Db 718 TCAAGTTCTAGTTATATGCAAAATCCAGTCAACCAAGATTGTCAGAGAACCAATCTG 777
QY 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
Db 778 ACTGTCATCCCACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAA 837
QY 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
Db 838 TTGTATGCTAAACCCCTTATCAGAACCCATGTTGAATCTGATGGCTTATTTTCGACCCA 897
QY 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
Db 898 GCGCAAAATCAAGTCGAACCGCCAGAGGTGTAGCTGTCCCTCATGTGGTAACCATTAACCAC 957
QY 301 PheIleProTyrGluGlnMetSerGluLeuGlnLysArgIleAlaArgIleIleProLeu 320
Db 958 TTTATCCCTTATGACAAATGTCTGAATGGAAAAACGAATTTGCTCGTATTATTTCCCTT 1017
QY 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGluProSerProGlnSer 340
Db 1018 CGTTATCGTTCAAAACCATTTGGGTACAGATTCAAGACAGAGAACCAAGTCCACACCG 1077

```



```
QY 341 ThrProGluProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
DB 1078 ACTCCAGAACCTAGTCCAACT-----CCGCAACCAAGCTCCAGCAAT 1119
QY 361 ProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
DB 1120 CCAATTGATGGGAATTTGGTCAAGAGAGCTGTCGAAAAAGTAGGCGATGGTATGTCTTT 1179
QY 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
DB 1180 GAGGAGATGGAGTTCTCTGTTATATCCAGCAAGGATCTTTCAGCAAGACAGCAGCA 1239
QY 401 GlyIleAspSerLysLeuAlaLysGlnGluSerLysHisLysLeuGlyAlaLysLys 420
DB 1240 GGCATTGATAGCAAACTGGCCCAAGCAGGAGAAAGTTTATCTCATAAAGCTAGGAACTAAGAAA 1299
QY 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
DB 1300 ACTGACCTCCCATCTAGTGCAGAAATTTTACAATAAGGCTTATGACTTACTTAGCAAGA 1359
QY 441 IleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
DB 1360 ATTACCAAGATTTACTTGATATAAAGTCCCAAGTTGATTTTGAGGCTTTGGATAAC 1419
QY 461 LeuLeuGluArgLysLeuAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
DB 1420 CTGTTGGAAACGACTCAAGGATGCTCAAGTGATATAAGTCAAGTACTGGAGATATTCIT 1479
QY 481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
DB 1480 GCCTTCTTGTAGTCCGATTCGTATCCAGAACGTTTGTAGAAAAACCAATGCGCAAAATTACC 1539
QY 501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
DB 1540 TACACTGATGATGAGTTCCAGTCCAGTTCGACGAGCAAGTACACAGCAAGACGGT 1599
QY 521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
DB 1600 TATATCTTTGATCCTCGTATATAACCACTGATGAGGGGATGCCCTATGTAACTCCACAT 1659
QY 541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAla 560
DB 1660 ATGACCCCATAGCCACTGGATTAAAAAGATAGTTGTCTGAAAGCTGAGAGCGCGAGCC 1719
QY 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
DB 1720 CAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCACAGACCATCAGGATTCAGGA 1779
QY 581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600
DB 1780 AATACTGAGGCAAAAGGAGCAGAGCTATCTACACCGMTGAAACGACTGAAGAGGTG 1839
QY 601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
DB 1840 CCACCTTGATCGTATGCTTACAACTTCAATATCTCAATATCTGTAAGAGTCAAAACCGTAGTTA 1899
QY 621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeu 640
DB 1900 ATCATACCTCATATTATGACCAATACCAATAACATCAAAATTTTGAAGTGGTTTGAAGAGCCCTT 1959
QY 641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrVal 660
DB 1960 TATGAGGCACCTPAGGGGTATATCTCTTGAGGATCTTTTGGGACGCTCAAGTACTATGTC 2019
QY 661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
DB 2020 GAACATCCAAACGAAAGCTCGCATTCAGATAATGTTTGGTAAACGCTAGCGACCATGTT 2079
QY 681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGluThrGluLysProSerGluGluLys 700
DB 2080 CAAGAGAAACAAATGGTCAAGCTGATACCAATCAACACGGAACCAAGCGAGGAGAAA 2139
QY 701 ProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGluLys 720
```

```
DB 2140 CCTCAGACAGAAAACCTGAGGAAGAAACCCCTCGAAGAGAAACCCCAAGCGAGAAA 2199
QY 721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
DB 2200 CAGAGTCTCAAAACCAACAGAGAACCAAGAGAAATCACCGAGGAATCAGAGAACT 2259
QY 741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuGlyLys 760
DB 2260 CAGGTCGAGACTGAAAGGTTGAAGAAAAAATCTGAGAGAGGCTGAAGATTTACTTCGAAA 2319
QY 761 IleGlnAsp 763
DB 2320 ATCCAGGAT 2328
RESULT 4
US-10-412-850-11
; Sequence 11, Application US/10412850
; Publication No. US20040001836A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-686
; CURRENT APPLICATION NUMBER: US/10/412,850
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 11
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-850-11
Alignment Scores:
Pred. No.: 0 Length: 2531
Score: 3919.00 Matches: 745
Percent Similarity: 98.03% Conservative: 3
Best Local Similarity: 97.64% Mismatches: 9
Query Match: 97.51% Indels: 6
DB: 16 Gaps: 1
US-09-765-272A-66 (1-763) x US-10-412-850-11 (1-2531)
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
DB 58 TGTTCCTATGAGCTTGAGCGTTACCAAGCTGTCAGGATAAGAAAGAGCTTAATCGAGTT 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
DB 118 GCTTATATAGATGGTGTATCAGGCTGGTCAAAAGCAGAAACCTTGACACCAAGTGAATC 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
DB 178 AGTAAGAGGAGGAGGATCAACGCGGAAACAAATGTTATCAAGATTACGATCAAGGTTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
DB 238 GTGACCTCTCATGAGACCAATATCATTTACTATTAATGCGAAGGTTCTCTATGATGCCATC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
DB 298 ATCAGTGAAGAGCTCTCTCATGAAGATCCGAATTTATCAGTTGAAGGATTCAGACATGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrValTyrLeuLys 120
DB 358 AATGAAATCAAGGCTGTTTGTGTCATTAAAGTAAACGCTAAATACTATGTTTACCTTAAG 417
```

121 AspAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
148 GATGCGGCTCATGCCGATATATTCGACAAAGAGAGAGATTAACCGTCAGAGCAGGAA 477
141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
478 CGCAGTCATATCATAACTCAAGAGCAGATATATGCTGTGTCAGCCAGAGCCCAAGGA 537
161 ArgTyrThrThraspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
538 CGTTATCAACGAGTATGATGCTATATCTCAATGCTATGATATCATTTGAGGACACGGGT 597
181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
598 GATGCTTATGCTTCTCAGCGGACCATTCATCATTAATCATTTCTAGATGATGATATCA 657
201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220
658 GCTAGCGAGTAGCTGCTGCAGAGCCTATTTGAATGGGAAGCAGGATCTCGTCCCTCT 717
221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
718 TCAAGTCTAGTTATTAATGCAATCCAGCTCAACCAAGATTGTTCAGAGAACCACAATCTG 777
241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
778 ACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTTACGTGAA 837
261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
838 TTGTATGCTTAACCTTATCAGAACCGCATGTGGAATCTGATGCGCTTATTTTCGCCCA 897
281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
898 GCGCAATCACAAGTCGAACCGCAGAGGTGTAGCTGCTCCCTCATGCTTAACCATACCA 957
301 PheIleProTyrGlnGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320
958 TTTATCCCTTATGAACAAATGTCTGAATTCGAAACAAAGAAATGCTGCTATTTATTTCC 1017
321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340
1018 CGTTATCGTTCAACCCATTGGTACCGATTCAGATTCAGACAGAGAACCAAGTCCACACCG 1077
341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
1078 ACTCCAGAACCTAGTCCAAGT-----CCGCAACCCAGCTCCCAAGCAAT 1119
361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
1120 CCAATTGATGGGAAATGGTCAAGAAGCTGTTCCGAAAGTAGCGATGCTATGCTCTTT 1179
381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
1180 GAGGAGAAATGGAGTTCTCGTTATATCCAGCCAGGATCTTTCCAGCAAGAACAGCAGCA 1239
401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420
1240 GGCATTCATAGCAAACTGGCCACAGCAAGATTTATCTCATAGCTAGGACTAGGAA 1299
421 ThrAspLeuProSerSerArgGluPheTyrAsnLysAlaTyrAspLeuAlaArg 440
1300 ACTGACCTCCATCTAGTATCGAGATTTTACAATAGGCTTATGATCTACTAGCAAGA 1359
441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
1360 ATTCAACCAAGATTTACTTGATATAAAGGTGCACAAAGTTGATTTGAGCGCTTGGATTAAC 1419
461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
1420 CTGTTGGAACGACTCAAGGATGTCTCAAGTGATAAAGTCAAGTAGTCAAGATATTTCTT 1479

481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
1480 GCCTTCTTAGCTCCGATTCGTCATCCAGACAGCTTTAGGAAACCAAAATGCGAAATTACC 1539
501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
1540 TACACTCATGATGATGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACAGCAAGACGGT 1599
521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
1600 TATATCTTTGATCTCGTGATATAACCAAGTATGAGGGGATGCTATGTAATCCACAT 1659
541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 560
1660 ATGACCCCATAGCACTGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGAGCC 1719
561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
1720 CAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCTCTCCACAGACCATCAGGATTCAGGA 1779
581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600
1780 AATCTGAGGCAAAAGAGAGCAGAACCTATCTACACCGMTGAACGAGCTAAGAGGTG 1839
601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
1840 CCACCTGATCGTATGCTTACAACTCTCAATATATCTGTAGAAGTCAAAACCGTAGTTTA 1899
621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640
1900 ATCATACCTCATATGACCATTAACCATTAACATCAAAATTTGAGTGGTTTGACGAAGCCCT 1959
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrVal 660
1960 TATGAGGCACCTAAGGGGTATCTCTTGAGGATCTTTTGCGGACTGTCAAGTACTATGTC 2019
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
2020 GAACATCCAAACCAAGCTCCGATTCAGATATGTTTGGTAACTAGCTAGCAGCAATGTT 2079
681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluLys 700
2080 CAAGCAAAACAAATGGTCAAGCTGATCAATCAAAACGCAAAACCAAGCGAGGAGAA 2139
701 ProGlnThrGluLysProGluGlnGluThrProArgGluGluLysProGlnSerGluLys 720
2140 CCTCAGACAGAAAAACCTGAGGAGAAACCCCTCGAAGAGAGAAACCGCAAGCGAGAA 2199
721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
2200 CCAGAGCTCTCCAAACCAACCAAGAGNACCAAGAGATCACCAGAGGANTCAGAGAACCT 2259
741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuLys 760
2260 CAGGTCGAGACTGAAAGGTTCAAGAAAAAATCTCAGAGAGCTCAAGATTTACTTGAATA 2319
761 IleGlnAsp 763
2320 ATCCAGGAT 2328
RESULT 5
US-10-387-783-11
; Sequence 11, Application US/10387783
; Publication NO. US20040005331A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13

; PRIOR APPLICATION NUMBER: 09/468,656
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/113,048
 ; PRIOR FILING DATE: 1998-12-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 11
 ; LENGTH: 2531
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-10-387-783-11

Alignment Scores:
 Pred. No.: 0 Length: 2531
 Score: 3919.00 Matches: 745
 Percent Similarity: 98.03% Conservative: 3
 Best Local Similarity: 97.64% Mismatches: 9
 Query Match: 97.51% Indels: 6
 DB: 16 Gaps: 1

US-09-765-272A-66 (1-763) x US-10-387-783-11 (1-2531)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGlySerAsnArgVal 20
 DB 58 TGTTCCTATGAGCTTGGAGCTTACCAAGCTGTCAGGATAAGAAAGAGCTTAATCGAGTT 117
 QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
 DB 118 GCTTATATAGATGGTGAATCAGGCTGCTCAAAAGGCGAGAAACTTGCACACAGATGAAGTC 177
 QY 41 SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr 60
 DB 178 AGTAAGAGGCGGGGATCAACCGCAACAAATGTTATCAAGATTACGATCAAGTTAT 237
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
 DB 238 GTCACCTCTCATGGAGACCATTAATCACTATAATGCGCAAGGTTCTTATGATGCCATC 297
 QY 81 IleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
 DB 298 ATCAGTGAAGAGCTCTCATGAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTGTC 357
 QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
 DB 358 AATGAATCAAGGCTGGTATGTCATTAAAGGTAACGGTAATACATATGTTACCTTAAG 417
 QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
 DB 418 GATGCGCTCATCGGATATATTCGGACAAAGAGAGATTAACCGTCAGACAGGAA 477
 QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly 160
 DB 478 CGCAGTCATAATCAATACTCAAGAGCAGATAATGCTGTGTCGAGCGAGCCCAAGGA 537
 QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
 DB 538 CGTTATACAACGGATGATGGGTATATCTTCAATGCACTGATATCATTTGAGGACACGGGT 597
 QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
 DB 598 GATGCTTATATCGTTCTTCACGGGACCAATTAACATTACATTTACATTCCTCAAGATGATCA 657
 QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220
 DB 658 GCTAGCGAGTTAGCTGCTGCGAGAGCCATTATGGAATGGCAAGCAGGATCTGTCCTCTCT 717
 QY 221 SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisLeuLeu 240
 DB 718 TCAAGTTCTAGTTATATGCAATCCAGCTCAACCAAGATTGTCAGAGAACACCAATCTG 777
 QY 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
 DB 778 ACTGTCACTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAA 837

QY 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
 DB 838 TTGTATGCTAAACCCCTTATCAGAACCCCATGTGGAATCTGATGCGCTTATTTTCGACCCA 897
 QY 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
 DB 898 GCGCAATCACAAGTCGAACCCGAGGCTGAGCTGCTCATGTGTAACCATTAACAC 957
 QY 301 PheIleProTyrGluGlnMetSerGluLeuGlyArgIleAlaArgIleIleProLeu 320
 DB 958 TTTATCCCTTATGACAAATGTCTGAATTCGAAACGAATTCGCTGATTTATTTCCCTT 1017
 QY 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340
 DB 1018 CGTTATCGTTCAAAACCATTTGGGTACAGATTCAAGCAGAGAACCAAGTCCCAACCG 1077
 QY 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
 DB 1078 ACTCCAGAACCTAGTCCAGT-----CGCAACACGAGCTCCAGCAAT 1119
 QY 361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
 DB 1120 CCAATTGATGGGAAATTTGGTCAAAAGAGCTGTTCCAAAAGTAGCGATGGTTATGCTTT 1179
 QY 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
 DB 1180 GAGGAGAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAGAAACACAGCA 1239
 QY 401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420
 DB 1240 GGCATTGATGCAAACTGGCCACAGCAGAAAGTTTATCTCATAGCTAGGAACCTAAGAA 1299
 QY 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
 DB 1300 ACTGACTCCCATCTAGTATCGAGATTTTACAATAAGGCTTATGACTTACTAGCAAGA 1359
 QY 441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
 DB 1360 ATTCAACAAGATTACTTGATAATAAGGTCGACAAAGTTGATTTTGAGGCTTTGGATAAC 1419
 QY 461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
 DB 1420 CTGTTGGAACACCTCAAGGATGTCCTCAAGTGATTAAGTCAAGTTAGTGAAGATATCTT 1479
 QY 481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
 DB 1480 GCCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAAAACCAATGCGCAAAATTACC 1539
 QY 501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
 DB 1540 TACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAGCAAGACGGT 1599
 QY 521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
 DB 1600 TATATCTTTGATCCCTCGTGATATAACCGATGATGAGGGGATGCTTATGATCTCCACAT 1659
 QY 541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAla 560
 DB 1660 ATGACCCATAGCCATCTGGATTAAAGAGATAGTTTCTGTAAGCTGAGAGCGGAGCC 1719
 QY 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
 DB 1720 CAGGCTTATGCTTAAAGAGAAAGGTTTACCCCTCTTCGACAGACCATCAGGATTCAGGA 1779
 QY 581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600
 DB 1780 AATACTAGGCAAAAGGAGCAGAAAGCTATCTACAACCGMGTAAGCAGCTAAGAGGTG 1839
 QY 601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
 DB 1840 CCACCTTGATCGTATGCTTCAATCTTCAATATACTGTAGAGTCAAAAACGGTAGTTTA 1899

621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeu 640
1900 ATCAATACCTCATTAAGACCATTTACCATTAACATCAATTTGAGTGGTTTGACGAGGCCTT 1959
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrVal 660
1960 TATGAGGCACCTAAGGGGTATACCTCTGAGGATCTTTGGCGACTGTCAAGTACTATGTC 2019
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
2020 GAACATCCAAACGACGTCGCAATTCAGATAATGGTTCGTAACGCTAGCGACCATGTT 2079
681 GluArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700
2080 CAAAGAAACAAATGCTGAGCTGATACCAATCAACGAGAAACCAAGCGAGGAGAAA 2139
701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720
2140 CCTCAGACAGAAACCTGAGGAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAAA 2199
721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
2200 CCAGAGTCTCCAAACCAACAGAGGAAACCAAGAGAAATCACCAGAGGAATCAGAGAACCT 2259
741 GlnValGluThrGluLysValGluGluLysArgGluAlaGluAspLeuGluLys 760
2260 CAGGTCGAGACTGAAAGGTTGAGAAACCTGAGAGAGGCTGAGATTACTTTGGAAA 2319
761 IleGlnAsp 763
2320 ATCCAGGAT 2328

RESULT 6

US-09-884-465A-5
; Sequence 5, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 2639
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-5

Alignment Scores:
Pred. No.: 0 Length: 2639
Score: 3665.00 Matches: 700
Percent Similarity: 92.78% Conservative: 20
Best Local Similarity: 90.21% Mismatches: 58
Query Match: 91.19% Indels: 18
DB: 10 Gaps: 3

US-09-765-272A-66 (1-763) x US-09-884-465A-5 (1-2639)

2y 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 171 TGTTCCTATGAACCTGGTGTGTCGCCACCAAGCTGGTCAAGTAAAGAAAGAGTCTAAATCGAGTT 230
2y 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40

Db 231 TCTTATATAGATGGTGAATCAGGCTGGTCAAAAGGCAGAGAAAATTTGACACCAAGATGAAGTC 250
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 291 AGTAAGAGAGAGGGGATCAACGCCGACAAATTTGTTATCAAGATTACGGATCAAGGTTAT 350
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 351 GTGACCTCTCATGGAGACCAATATCATTTACTATAATGCGAGGTTCTTTATGATGCATC 410
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 411 ATCAGTGAAGAACTCTCATGAAGATCCGAATATCAGTTTGAAGGATTCAGACATGTC 470
Qy 101 AsnGluLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 471 AATGAATCAAGGTTGGCTGTGTTAAGGTAGACGGAATACTATGTTTACCTTAA 530
Qy 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 531 GATCGGCCCATCGGACAATATTCGACAAAGAGAGATTAAACGTCAGAAAGCAGGAA 590
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaAlaGlnGly 160
Db 591 CACAGTCAATAATCAATCACTCAAGACGAGATAATGCTTGTCTGCACGACGAGCCCAAGGA 650
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 651 CGTTATACACGATGATGGGTATATCTTCATGCAATCTGTATATCATTTGAGGACACGGGT 710
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 711 GATGCTTATATCGTTCTCCAGCGGACCATTAACATTACATTCCTAAGAATGAGTTATCA 770
Qy 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220
Db 771 GCTAGCGAGTTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGGATCTCGTCTCTCT 830
Qy 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
Db 831 TCAAGTTCTAGTTATTAATGCAATCCAGTTCACCAAGATTGTCAAGAAACCAATCTG 890
Qy 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260
Db 891 ACTGTCACTCCCACTTAUATCAATAATCAAGGGGAAACATTTCAAGCCTTTTACGTA 950
Qy 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280
Db 951 TTGTATGCTAAACCTTATCAGAACGCCATGTAGAAATCTGATGGCCTTATTTTCGACCCA 1010
Qy 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
Db 1011 GCGCAATCAAGTCGAACCCGACGAGGTGAGCTGTCCCTCATGGTAACCATACCAC 1070
Qy 301 PheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320
Db 1071 TTTATCCCTTATGAACAAATGTCTGAATTTGGAAAACGAATTCCTCTGCTATTATTCCTCT 1130
Qy 321 ArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSer 340
Db 1131 CGTTATCGTTCAACCCATTGGGTACCAAGATTCAAGACCAAGAAACCAAGTCCCAATCG 1190
Qy 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
Db 1191 ACTCGGAACCTAGTCCAAAGTCTGCAACCTGCAACCAATCTCAACGAGTCCCAAGCAAT 1250
Qy 361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
Db 1251 CCAATTTGATGAGAAATTTGTCGAAGAGCTGTTGAAAAGTAGGCGATGTTATGCTTT 1310
Qy 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
Db 1311 GAGGAGATGAGGTTTCTCGTTATATCCCGCAAGGATCTTTTTCAGCAGAAACAGCAGCA 1370

QY 401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyValAlaLysLys 420
Db 1371 GGCATTGATAGCAAACTGGCCAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAAA 1430
QY 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
Db 1431 ACTGACCTCCCATCTAGTATCGAGATTTTACAAATAAGGCTTATGACTTATGACGAGA 1490
QY 441 IleHisGlnAspLeuLeuAspAsnLysGlyValArgGlnValAspPheGluAlaLeuAspAsn 460
Db 1491 ATTCAACAAGATTACTTGATAATAAAGGTCGCAAGTTGATTTGATTTGGATTAAC 1550
QY 461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspileLeu 480
Db 1551 CTGTGGAAAGCACTCAAGGATGCTCAAGTATTAAGTCAAGTTAGTGATGATATTCTT 1610
QY 481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
Db 1611 GCCTTCTTAGCTCCGATTCGTATCCAGAACGTTTAGGAAACCAATGCGCAAAATTACC 1670
QY 501 TyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
Db 1671 TACACTGATGATGATGATCAAGTATGAGCAAGTTGGCAGCAAGTACACCAAGAACCGGT 1730
QY 521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
Db 1731 TATATCTTTGATCTCGTGATATTAACCAAGTATGAGGGGATGCTATGTAACCTCCACAT 1790
QY 541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 560
Db 1791 ATGACCCATAGCCACTGGATTGATTAAGATAGTTTGTCTGAAGCTGAGAGCGCGCAGCC 1850
QY 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
Db 1851 CAGGCTTATGCTAAAGAAAGGTTTGACCCCTCCCTCCGACAGACCCACGAGATTACAGA 1910
QY 581 AsnThrGluAlaLysGlyAlaGluAlaLysTyrAsnArgValLysAlaAlaLysLysVal 600
Db 1911 AATACTGAGGCAAAAGGAGCAGAGCTATCTACAAACCGGTGAAGCAGCTAAGAGGTG 1970
QY 601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
Db 1971 CCACCTGATGATGCTTCAAACTCTCAATATATCTGTAGAGTCAAAACCGGTAGTTTA 2030
QY 621 IleIleProHisTyrAspHisTyrHisAsnLysPheGluTyrPheAspGluGlyLeu 640
Db 2031 ATCATACCTCATATTATGACCATTAACCATTAACATTAATTTGATGTTTGACAGGCGCTT 2090
QY 641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrVal 660
Db 2091 TATGAGGCACTTAAGGGGTATAGTCTTGAAGATCTTTGGCAGCTGTCAAGTACTATGTC 2150
QY 661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
Db 2151 GAACATCCAAACGACGCTCCGATTCAGATAATGTTTGGTAAACGCTAGTCAACATGTT 2210
QY 681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700
Db 2211 CGTAAAAATAAG-----GCAGACCAAGATAGTAAACCTGATGAAGATAAG 2255
QY 701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720
Db 2256 GAACATGATGAAGTAAAGTGAAGCACTACCTCAATCTGATGATAAAGAGAAATCAGCT 2315
QY 721 ProGluSerPro-----LysProThrGluGluProGluGluSerPro 734
Db 2316 GGTTTAAATCTTACGACATAATCTTTATAAACCAAGCACTGATACGGAAGACAGAG 2375
QY 735 GluGluSerGluGlu-----ProGlnValGluThrGluLysVal 747
Db 2376 GAAAGAGCTGAAGATACCAAGATGAGGCTGAAATTCCTCAAGTAGAAGATTCGTATT 2435

QY 748 GluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763
Db 2436 AACGCTAAGATAGCAGATCGGAGCGCTTGCTAGAAAAAGTAGAACAGAT 2483
RESULT 7
US-10-412-862-5
; Sequence 5, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-862-5
Alignment Scores:
Pred. No.: 0 Length: 2531
Score: 3644.50 Matches: 698
Percent Similarity: 92.41% Conservative: 20
Best Local Similarity: 89.83% Mismatches: 40
Query Match: 90.68% Indels: 19
Gaps: 4
US-09-765-272A-66 (1-763) x US-10-412-862-5 (1-2531)
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 58 TGTTCCTATGAATCTGCTGCTCACCAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTT 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATATAGTGTGTCATCAGGCTGCTCAAAAGCGCAGAAAACCTGACACAGATGAAGTC 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGTAAGAGGGAGGGGATCAACGCCGAACAAATCGTCATCAAGATTACGGATCAAGGTTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTGACCTCTCATGAGACCACTTATCTACTATAATGCGAAGGTCCTTATGATGCCATC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 ATCAGTGAAGAGCTCTCTCATGAAGATCCGAATATTATCAGTTGAAGGATTTCAGACATTGTC 357
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrIleLys 120
Db 358 AATGAAATCAAGGCTGCTTATGTTATCAAGGTAGATGGAATAATACTATGTTTACTCTTAAG 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGCGAGCTCATCGGATAATATTCGGACAAAGAGAGATTAAACGTCAGACGAGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAraGln 159
Db 478 CACAGTCATTAATCACGGGGGTGGTCTTAACGATCAAGCAGTAGTTGCGAGCCAGAGCCAA 537
QY 160 GlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr 179

Alignment Scores:

Pred. No.: 0 Length: 2531
 Score: 3644.50 Matches: 698
 Percent Similarity: 92.41% Conservative: 20
 Best Local Similarity: 89.83% Mismatches: 40
 Query Match: 90.68% Indels: 19
 DB: 16 Gaps: 4

US-09-765-272a-66 (1-763) x US-10-412-850-5 (1-2531)

QY	1	CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal	20	958	CAC	TTATCCCTTATGACAAATGCTCTGAATTTGAAAAACGAATTTGCTCGTATTATTATCC	1017
DB	58	TGTTCTCTATGAACTTGGTGGTCCACCAAGCTGTCAGGTGAAGAAAGAGTCTAATCCAGTT	117	320	LeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGln	339	1018
QY	21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40	1018	CTTCGTTATCGTTCAAAACCATTTGGGTACCAAGATTTCAAGACCAGAACCAACCAAGTCCACAA	1077	340
DB	118	TCTTATATAGATGGTGATCAGCGCTGGTCAAAAGCGCAAAATCTTGACACAGATGAAGTC	177	1078	TCGACTCCGGAACCTAGTCCAAAGTCCGCAACCTCGCAACAAATCTCTCAACCCAGCTCCAAAGC	1137	1078
QY	41	SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr	60	360	AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal	379	1138
DB	178	AGTAAGAGGAGGGATCAACCGCAACAAATCGTCAACAGATTACGGATCAAGGTAT	237	1138	AAATCCCAATTCATGAGAAATTTGGTCAAAAGAGCTGTTGAAAAGTAGGCGATGGTTATGTC	1197	380
QY	61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80	380	PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla	399	1198
DB	238	GTGACCTCTCATGAGACCAATTCATTACTATAATGCAAGGTCCCTTATGATGCCATC	297	1198	TTTGAGGAGAAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTCCAGCAGAACAGCA	1257	400
QY	81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysAspSerAspIleVal	100	1258	GCAGGCATTCATAGCAAACTGGCCAGCAGCAAGAAATTTATCTCATAGCTAGGAGCTAAG	1317	420
DB	298	ATCAGTGAAGACTCTCTATGAAGATCCGAATATCAGTTGAAGATTCAGACATGTC	357	1318	AAAACCTGACCTCCCATCTAGTATCGAGAAATTTTACAAATAAGGCTTATGACTTACTAGCA	1377	440
QY	101	AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120	440	ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp	459	1378
DB	358	AATGAATCAAGGGTGTATGTATATCAAGGTAGATGGAATACTATGTTTACCTTAAG	417	1437	AGAAATTCACCAAGATTTACTTTGATATAAGTTCGACNAGTTGATTTTGGCTTTGGAT	1437	460
QY	121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140	1438	AACTGTTTGAAGACACTCAAGGATGTCCTCAAGTGTATAAAGTCAAGTTAGTGGATGATATT	1497	480
DB	418	GATGACGCTCATCGCGATATATTCGCAAAAGAGAGATTAACAGCTCAGAAAGCAGAA	477	1498	CTTGCCCTCTTAGCTCCGATTCGTCATCCAGAACTTTAGGAAACCAAAATGCGCAAAAT	1557	500
QY	141	ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAlaArgAlaGln	159	1558	ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp	519	1617
DB	478	CACAGTCATAATCAACGGGGTGGTCTTAACGATCAAGCAGTAGTTCGACGACAGGCCAA	537	1617	ACCTACACTGATGATGAGATTCAGTAGCCNAGTTGGCAGGCAAGTACACACAGNAGAC	1617	520
QY	160	GlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr	179	1618	GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro	539	1677
DB	538	CGACGCTATACACGATGATGTTATATCTTCAATGCACTGATTCATTTGAGGACAG	597	540	HiMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla	559	1678
QY	180	GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu	199	1678	CATATGACCCATAGCCACTGGATTTAAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCA	1737	560
DB	598	GGTGATGCTTATATCGTCTCCACGCGCACCAATTACCATTTACATTCCTAAGATGAGTTA	657	1738	GCCAGGCTTATGCTAAAGAGAAAGTTTGACCCCTCTTCACACAGCATCAGGATTC	1797	580
QY	200	SerAlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgPro	219	1798	GGAATACTAGGCAAAAGAGGACAGAGCTATCTACACCCGCGTGAAGACAGCTAAGAG	1857	600
DB	658	TCAGCTAGCGAGTAGTCTGCTCGAAGGCTATTTGAATGGGAAGCAGGGATCTCGTCT	717	600	ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer	619	1858
QY	220	SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn	239	1858	GTGCACTTATGCTATGCTCTTACATCTTCAATATATCTGTAGAAGTCAAAAACCGGTAGT	1917	620
DB	718	TCTTCAGTCTTAGTTATTAATGCAATTCAGCTCAACCAAGATTCAGAGAACCAAT	777	620	LeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGly	639	1918
QY	240	LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg	259	640	LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr	659	1978
DB	778	CTGACTGTCTCACTCCAACTTATCATCAAAATCAAGGGGMAAAACATTTCAAGCCCTTTAGCT	837	660	ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis	679	2038
QY	260	GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp	279	2038	GTGAAACATCCAAACGAACGTCGCCATTCAGATAATGTTTTGTTGTAACGCTAGCGACCAT	2097	
DB	838	GAATTTGATGCTAAACCCCTTATCAGAACGCCATGTGGAATCTGATGGCCCTTATTTTCGAC	897				
QY	280	ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr	299				
DB	898	CCAGCGCAATTCACAGTCCAAACCGCCAGAGGTAGTGTCTCCCTCATGTGTAAACCATTC	957				
QY	300	HisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIlePro	319				

Db 1378 AGAATTCACCAAGATTACTGTGATAATAAAGCTCGACAAGTTGATTTTGGAGCTTTGGAT 1437
Qy 460 AsnLeuLeuGluArgLeuLysAspVal***SerAspLysVallyLysLeuVal***AspIle 479
Db 1438 AACCTGTTGGACGACTCAAGGATGTCCTCAAGTGTATTAAGTCAAGTTAGTGTGATGATTT 1497
Qy 480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499
Db 1498 CTTGGCTCTTCTAGCTCGATTGCTGATTCACAGACGTTTGTAGGAAACCAAAATGGCCAAAT 1557
Qy 500 ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp 519
Db 1558 ACCTACACTGATGATGATGATCAAGTACCAAGTTGGCAGGCAAGTACACACAGAGAC 1617
Qy 520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro 539
Db 1618 GGTATTATCTTTGATCTCTGATATACCAAGTATGAGGGGATGCTTGTGTAATCTCA 1677
Qy 540 HisMetThrHisSerHisTrrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 559
Db 1678 CATATGACCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGACGGCA 1737
Qy 560 AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThrAspHisGlnAspSer 579
Db 1738 GCCCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTC 1797
Qy 580 GlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVallyLysAlaLysLys 599
Db 1798 GGAATACTGAGGCAAAAGGAGAGAGACTATCTACAACCGGTGAAAGCAGCTAAGAG 1857
Qy 600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619
Db 1858 GTGCCACTTGATCGTATGCTTACATCTTCAATATCTGTAGAGTCAAAAACGGTAGT 1917
Qy 620 LeuIleleProHisTyrAspHisTyrHisIleLysPheGluTyrPheAspGluGly 639
Db 1918 TTAATCATACCTCATATGACCAATACCATAACATCAAAATTTGAGTGGTTTGACGAAGC 1977
Qy 640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrVallystTyrTyr 659
Db 1978 CTTTATGAGGCACCTTAAGGGGTATCTCTTGAGGATCTTTTGGGACTGTCAGTACTAT 2037
Qy 660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679
Db 2038 GTCCACATCAACAGACGCTCCGATTCAGATATGTTTGGTAAACGCTAGCAGCAT 2097
Qy 680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699
Db 2098 GTTCGTAATAATAAG-----GTAGACCAAGACAGACAGTAAACCTGATGAAGAT 2142
Qy 700 LysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGlu 719
Db 2143 AAGGAACATGATGATGATGAGTAAAGTGAAGTCACTACCTCGAATCTGATGAAAGAGATC 2202
Qy 720 LysProGluSerPro-----LysProThrGluGluProGluGluSer 733
Db 2203 CTGCTTTAAATCTTCACGAGATTAATCTTTATAACCAAGCACTGATACGGAAGAGACA 2262
Qy 734 ProGluGluSerGluGlu-----ProGlnValGluThrGluLys 746
Db 2263 GAGGAAGAAGCTGAAGATACCAACAGATGAGGCTGAAATTCCTCAAGTACAGAAATCTGTT 2322
Qy 747 ValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763
Db 2323 ATTAACGCTAAGATAGACAGATCGGAGGCTTGCTAGAGAAAAATACAGAT 2373

RESULT 10

US-10-158-844-94

; Sequence 94, Application US/10158844

; Publication No. US20040029118A1

; GENERAL INFORMATION:

; APPLICANT: Kunsch et al.

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31

; ATTORNEY/AGENT INFORMATION:

; NAME: Hyman, Mark J.

; REGISTRATION NUMBER: 46,789

; REFERENCE/DOCKET NUMBER: PB340P1D1

; INFORMATION FOR SEQ ID NO: 94:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8195 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 94:

US-10-158-844-94

Alignment Scores:

Pred. No.: 0 Length: 8195

Score: 3641.50 Matches: 698

Percent Similarity: 92.41% Conservative: 20

Best Local Similarity: 89.83% Mismatches: 40

Query Match: 90.61% Indels: 19

DB: 13 Gaps: 4

US-09-765-272A-66 (1-763) x US-10-158-844-94 (1-8195)

Qy	1	CysSerTyrGluLeuGlyArgHisGlnAlaGlyLysGluSerAsnArgVal	20
Db	5659	TGTTCTCTATGAATGGTCTGCCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTT	5718
Qy	21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
Db	5719	KCTTATATAGATGGTGCATCAGGCTGGTCAAAAGGACGAAAACTTGACACAGATGAAGTC	5778
Qy	41	SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr	60
Db	5779	AGTAAGAGGAGGGGATCAACGCCGAACAAATCGTCATCAAGATTACGGATCAAGGTTAT	5838
Qy	61	ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle	80
Db	5839	GTGACCTCTCATGGACCACTTATCTATTAATGCGAAGGTCCTTATGATGCCATC	5898
Qy	81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal	100
Db	5899	ATCAGTGAGAGAGCTCTCATGAAGATCCGAATATCAGTTGAAGATTCAGACATTGTC	5958
Qy	101	AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLysLys	120
Db	5959	AATGAAATCAAGGCTGGTTATGTTATCAAGGTAGATCGAAAAATACTATGTTTACTTAAG	6018
Qy	121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluLysArgGlnLysGluGlu	140
Db	6019	GATGCACTCATCGGATATATATTCGACACAAAAGAGATTAAACCTCGAAGACGAA	6078

141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAlaArgAlaGln 159
6079 CACAGTCATAATCAGGGGTGGTCTTAACGATCAAGCGAGTAGTTCAGCCAGAGCCCAA 6138
160 GlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleLeuGluAspThr 179
6139 GGACGCTATACACGGATGATGGTTATATCTTCAATGTCATCTGATATCATTTAGGACACG 6198
180 GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu 199
6199 GGTGATGCTTATATCTTCTCAGCGGACCAATACCATTTACATTCCTAAGAATGAGTTA 6258
200 SerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgPro 219
6259 TCAGCTAGCGAGTTAGCTGCTCGAAGAGCCCTATGAAATGGGAAGCAGGATCTCGTCT 6318
220 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn 239
6319 TCTTCAAGTCTAGTTATATGCAATCCAGCTCAACCAAGATGTTCAGAGAACCACAAT 6378
240 LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259
6379 CTGACTGTCTACTCAACTTATCATCAAAATCAAGGGGAAAACATTTCAAGCCTTTTACGT 6438
260 GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279
6439 GAATGTGATGCTAAACCTTATCAGACGCGCATGTGGAATCTGATGCGCTTATTTTCAC 6498
280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299
6499 CAGCGCAATACACAGTCAACCGCAGAGGTAGTGTCTCCCTCATGTTAACCATTAC 6558
300 HisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIlePro 319
6559 CACTTATCCCTTATGAACAAATGCTGAAATGGAATAACAAATGCTGCTGATTTATTC 6618
320 LeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGln 339
6619 CTTTCGTATCTGTTCAACCAATGGGTACCAAGATTCAAGACCAGAACCAACCAAGTCCCAA 6678
340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359
6679 TCAGCTCCGGAACCTAGTCCAAAGTCCGCAACCTGCACCAATCTCAACAGCTCCAGC 6738
360 AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal 379
6739 AATCCATTCATGAGAAATGGTCAAGAAGCTGTTCGAAAAGTAGCGATGGTTATGTC 6798
380 PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla 399
6799 TTTGAGGAGAATGGAGTTTCTCGTTATATCCAGCAAGGATCTTTCAGCAAGAACACGA 6858
400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys 419
6859 GCAGGCATTTAGTACAACTGGCCAGCAGCAAGAAAGTTATCTCATAGCTAGGAGCTAAG 6918
420 LysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAla 439
6919 AAACTGACCTCCCATCTAGTATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGCA 6978
440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp 459
6979 AGAATTCACCAAGATTTACTTGTATTAATAAAGTCGACCAAGTTGATTTTGAAGCTTTGGAT 7038
460 AsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIle 479
7039 AACCTGTTGGAAACGACTCAAGATGTCYCAAGTGATAAAGTCAAGTTAGTGTGATGATAT 7098
480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499
7099 CTTGCTTCTTGTAGCTCCCATCTGTCATTCAGAACGTTTAGGAAACCAATCGGCAAT 7158
500 ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAsp 519

7159 ACCTACTACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAACAGAGAC 7218
520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro 539
7219 GGTATATCTTTGATCTCTCGATATACCAAGTATGAGGGGATGCCATCTATCTAATCCA 7278
540 HisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 559
7279 CATATGACCCCATAGCCACTGGATTAAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCA 7338
560 AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer 579
7339 GCCCAGGCTTATGCTATAGAGAAAGGTTTGAACCTCTCTTCACAGACCATCAGGATTTCA 7398
580 GlyAsnThrGluAlaLysGlyValaGluAlaIleTyrAsnArgValLysAlaAlaLysLys 599
7399 GGAATATCTGAGGCAAAAGGAGCAGAGCTATCTACACCGCGTGAAGCAGCAGTAAGAAG 7458
600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619
7459 GTGCCACTTGTATCGTATGCTTACATCTTCAATATCTGTAGAACTCAAAAACGCTAGT 7518
620 LeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluThrPheAspGluGly 639
7519 TTAATCATACCTCATATGACCATTAACCATCAATCAATTTGAGTGGTGTGACGAAGGC 7578
640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr 659
7579 CTTTATGAGGACCACTTAAGGGGTATATCTTGTAGGATCTTTTGGCGACTGTCAAGTACTAT 7638
660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679
7639 GTCGAACATCCAAACGACGTCGCAATTCAGATAATGGTTTGTAAACGCTAGCAGCAT 7698
680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699
7699 GTTCTGTAATAATAAG-----GTAGACCAACAGACAGTAAACCTGATGAAGAT 7743
700 LysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGlu 719
7744 AAGGAACATGATGAAGTAAAGTAGTGAGCAACTCACCTGAAATCTCATGAAAAAGAGAAATCAC 7803
720 LysProGluSerPro-----LysProThrGluGluProGluGluSer 733
7804 GCTGGTTAAATTCCTTCAGCAGATAATCTTTATAACCAAGCACTGATACGGAAGACACA 7863
734 ProGluGluSerGluGlu-----ProGlnValGluThrGluLys 746
7864 GAGGAAGAAGCTGAAGATACCAAGATGAGGCTGAAATCTCTCAAGTAGAGATTTCTGTT 7923
747 ValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763
7924 ATTAAGCTTAAGATAGCAGATGCGGAGGCTTCTGTAAGAAAAAGTAAACAGAT 7974

RESULT 11

US-09-884-465A-3
; Sequence 3, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884, 465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-3

Alignment Scores:
  Pred. No.: 3,63e-315      Length: 2523
  Score: 3588.50           Matches: 693
  Percent Similarity: 91.68% Conservative: 23
  Best Local Similarity: 88.73% Mismatches: 44
  Query Match: 89.29%      Indels: 21
  DB: 10                   Gaps: 6

US-09-765-272A-66 (1-763) x US-09-884-465A-3 (1-2523)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
D 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 58 TGTGCTTATGAAGTATGTTTGCATCAAGCT---CAAACCTGTAAAGAAAATAATCGTGT 114
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
D 115 TCCTATATAGTGAAGAAACAAGCGACGCAAAAACGGAGAAATTTGACTCTCCTGATGAGGTT 174
QY 41 SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr 60
D 175 AGCAGCGGTGAAGGAATCAACGCCGACAAATCGTCATCAAGATTACGGATCAAGTTAT 234
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
D 235 GTGACCTCTCATGGAGACCAATATATCTATCAATATGCAAGGTCCTTATGATGCCATC 294
QY 81 IleSerGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
D 295 ATCAGTGAAGAGCTCCTCATGAAGATCCGAATATATCATGTTGAAGGATTTCAGACATTGTC 354
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
D 355 AATGAATCAAGGCTGTTATGTCTATTAAGTAAACGGTAATACTATGTTTACCTTAAG 414
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
D 415 GATGCAGCTCATGGCGGATAATCGTCAAAAGAAAGAAATCAATCGGCAAAAACAAGAA 474
QY 141 ArgSerHisAsnHis-----AsnSerArgAlaAspAsnAlaValAlaAlaAlaArg 157
D 475 CATAGTCAGCATCGTGAAGGAGGACTTCAGCAACGATGTGGGTAGCCTTTGCACGT 534
QY 158 AlaGlnGlyArgTyrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177
D 535 TCACAGGAGCGCTACACACAGATGATGTTATATCTTCAATGCATCTGATATCATCGAA 594
QY 178 AspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsn 197
D 595 GATACGGCGCATGCTATATCGTCTCATGAGATCATATACCATATACCATTCCTAAGAA 654
QY 198 GluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGly--- 216
D 655 GAGTTATCAGTAGCGAGTGTGGCTGCTCAAGGCTTCCTATCTGTGGTGGGAAAATCTG 714
QY 217 -----SerArgProSerSerSerSerSerSerTyrAsn----- 226
D 715 TCAAAATTAAGAACTATCGCGGACAAAATAGCGNTAACTCCAGAACAAACTGGGTA 774
QY 227 -----AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThr 243
D 775 CTTCTCTGTAAAGCAATCCAGGAATACAAATACTAACAACAAGCAACAACAAGCAACTAAC 834
QY 244 ProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAla 263
D 835 AGT-----CAAGCAAGTCAAAAGTAATGACATTTGATAGTCTCTTGAAGACAGCTCTCAAA 888
264 LysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIle 283
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
889 CTGCTTTTCAGTCAAGCCCATGTAGTAATCTGATGGCTTATTTTCGACCCAGGCAAAATC 948
QY 284 ThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIlePro 303
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
949 ACAAGTCGAACCCGACAGAGGTGAGCTGTCCCTCATGGTAACCATTACCATTATCCCT 1008
QY 304 TyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArg 323
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1009 TATGAACAATGTCGAATTTGAAAACGAATTTGCTCGTATTATTCCTCCCTTCGTATCGT 1068
QY 324 SerAsnHisTyrPheValProAspSerArgProGluGlnProSerProGlnSerThrProGlu 343
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1069 TCAAAACCATTTGGGTACAGATTCAAGCAGAGAACAAGTCCCAACACGACTCCAGAA 1128
QY 344 ProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAsp 363
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1129 CCTAGTCCCAAGTCCGCAACCTGCACCAATCTTCAACAGCTCCAGCAATCCAAATTTGAT 1188
QY 364 GluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsn 383
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1189 GAGAAATTTGGTCAAGAAGCTGTTGAAAAGTAGGCGATGTTATGCTCTTTGAGGAAAT 1248
QY 384 GlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAsp 403
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1249 CGAGTTTCTCGTTATATCCAGCCAAAGATCTTTCAGCAAAACAGCAGCAGGCAATTTGAT 1308
QY 404 SerLysLeuAlaLysGlnLysLeuSerHisLysLeuGlyAlaLysLysThrAspLeu 423
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1309 AGCAAACTCGCCCAAGCAGGAAATTTATCTCATAGCTAGAGCTAAGAAAACCTGACCTC 1368
QY 424 ProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGln 443
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1369 CCATCTAGTGTGCGAATTTTACAATTAAGCTTATGACTTACTAGCAGAATTCACAA 1428
QY 444 AspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGlu 463
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1429 GATTTACTTGATAATAAAGTGCACAAGTTGATTTTGGGCTTTGGATAACCTGTTCGAA 1488
QY 464 ArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeu 483
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1489 CGACTCAAGGATGCTCTCAAGTGTATAAAGTCAAGTTAGTGTGATGATATTTCTGCTTCTTA 1548
QY 484 AlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAsp 503
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1549 GCTCCGATTCGTCATCCCAAGCTTTAGGAAAACCAAAATGCGCAATTTACTACACTGAT 1608
QY 504 AspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePhe 523
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1609 GATGAGATTCAGTAGCCCAAGTTGGCAGCAAGTACACACAGACAGCGGTATATCTTT 1668
QY 524 AspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHis 543
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1669 GATCTCTCGTATATAACCAAGTGTAGGGGGGATGCTATGTAATCCACATATGACCCAT 1728
QY 544 SerHisTyrPheLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyr 563
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1729 AGCCACTGATGATTAAGAAAGATAGTTGTCTGAAGCTGAGAGAGCGGACGCCAGGCTTAT 1788
QY 564 AlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGlu 583
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1789 GCTAAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTCAGGAAATACTGAG 1848
QY 584 AlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAsp 603
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1849 GCAAAAGGAGCAGAGAGCTATCTACACCGCTGAAAAGCAGCTAAGAAGGTGCCACTTGTAT 1908
QY 604 ArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIlePro 623
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1909 CGTATGCTCTACAACTTCAATATATCTAGTAAGTCAAAACGCTAGTTTAAATCATACCT 1968
QY 624 HisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAla 643
```

1969 CATTATGACCATTAACATAACATCAAAATTGAGTGGTTTGAAGAGGCCTTTATGAGGCA 2028
 644 ProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisPro 663
 2029 CCTAAGGGGTATACCTCTTGGAGGATCTTTGGCGACTGTCAAGTACTATGTGGAACATCCA 2088
 664 AsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsn 683
 2089 AACGNAAGTCCGCAATTCAGATTAATGGTTTGGTAAGCTAGCGACCATGTTCAGAGAAC 2148
 684 LysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThr 703
 2149 AAAAATGGTCAAGCTGATACCAATCAACCGGAAAAACCAAGCGAGGAGAACCTCAGACA 2208
 704 GluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSer 723
 2209 GAAAAACCTGAGGAGAAACCCCTCGAGAAGAGAAACCAAGAGCGAGAACACAGAGTCT 2268
 724 ProLysProThrGluGluPro--GluGluSerProGluGluSerGluGluProGlnVal 742
 2269 CAAAACCAACAGAGAACCAAGAGAGAAATCAACAGAGGAATCAGAGAACCTCAGGTC 2328
 743 GluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGln 762
 2329 GAGACTGAAAAGGTTGAAGAAAACTGAGAGAGGCTGAGATTCTTGGAAAAATCCAG 2388
 763 Asp 763
 2389 GAT 2391

RESULT 12

US-09-884-465A-4
 Sequence 4, Application US/09884465A
 Publication No. US2003007293A1

GENERAL INFORMATION:

APPLICANT: Shire Biochem, Inc.
 APPLICANT: Hamel, Josee
 APPLICANT: Brodeur, Bernard
 APPLICANT: Martin, Denis
 APPLICANT: Charland, Nathalie
 APPLICANT: Ouellet, Catherine
 TITLE OF INVENTION: Streptococcus Antigens
 FILE REFERENCE: 055190-0044
 CURRENT APPLICATION NUMBER: US/09/884,465A
 CURRENT FILING DATE: 2001-06-20
 PRIOR APPLICATION NUMBER: 60/212,683
 PRIOR FILING DATE: 2000-06-20
 NUMBER OF SEQ ID NOS: 384
 SOFTWARE: Patent version 3.1
 SEQ ID NO 4
 LENGTH: 2647
 TYPE: DNA
 ORGANISM: Streptococcus pneumoniae

JS-09-884-465A-4

Alignment Scores:

Score: 3,896-315 Length: 2647
 Pred. No.: 3588.50 Matches: 693
 Percent Similarity: 91.68% Conservative: 23
 Best Local Similarity: 88.73% Mismatches: 44
 Query Match: 89.29% Indels: 21
 Gaps: 6

JS-09-765-272A-66 (1-763) x US-09-884-465A-4 (1-2647)

2Y 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
 102 TGTGCTTATGAACCTAGCTTTGCTGCTCAAGCT---CAAACTGTAAAGAAAAATAATCGTGT 158
 2Y 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
 159 TCCTATATAGTGGAAACAAAGCGACGCAAAACCGAGAAATTTGACTCTCTGAGGTT 218

QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
 DB 219 AGCAAGCGTGAAGGAATCAACGCCGAACAAATCGTCATCAAGATTACGATCAAGGTTAT 278
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
 DB 279 GTGACCTCTCATGGAGACCATTAATCTACTATATATGCAAGGTCCTTATGATGCCATC 338
 QY 81 IleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
 DB 339 ATCAGTGAAGAGCTCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCACAGATTGTC 398
 QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLysLys 120
 DB 399 AATGAATCAAGCGTGGTTATGTCAATTAAGGTAAACGGTAAATCTATGTTTACCTTAAG 458
 QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
 DB 459 GATCAGCTCATGCGGATAATGTCCGTACAAAGAGAAATCAATCGGCAAAACAAAGAA 518
 QY 141 ArgSerHisAsnHis-----AsnSerArgAlaAspAsnAlaValAlaAlaArg 157
 DB 519 CATAGTCAGCATCTGTGAAGGAGGACTTCAGAAACGATGGTGGTAGCCTTGCACCT 578
 QY 158 AlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177
 DB 579 TCACAGGAGCGCTACACACAGATGATGTTTATATCTTCAATGTCATCTGATATCATGAA 638
 QY 178 AspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsn 197
 DB 639 GATACGGGCGATGCTATATGCTTCTCATGAGAGATCATTTACCATTTACATTTCTTAAGAAT 698
 QY 198 GluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGlnGly--- 216
 DB 699 GAGTTATCAGCTAGCTAGTGGCTGCTGCAAGGCTTCTCTATCTGCTGGGAAATCTG 758
 QY 217 -----SerArgProSerSerSerSerSerSerSerSerSerSerSerSerSerSer 226
 DB 759 TCAATTTAAGAACTTATCGCGCAGCAAAATAGCGATACACTCCAAGAACAACTGGGTA 818
 QY 227 -----AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThr 243
 DB 819 CTTTCTGTGAAGCAATCCAGGAATACAAATACTAAACAGCAACCAACCAACCACTAAC 878
 QY 244 ProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAla 263
 DB 879 AGT-----CAAGCAAGTCAAAAGTAATGACATGATAGTCTCTTCAAAACAGCTCTACAA 932
 QY 264 LysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIle 283
 DB 933 CTGCTTTGAGTCAAGCCATGTAGAACTGATGGCCCTTATTTTCGCCCGCGCAATC 992
 QY 284 ThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIlePro 303
 DB 993 ACAAGTCGAACCGCAGAGGTGTAGCTGTCCCTCATGTTAAACCATTTACCATTTATCCCT 1052
 QY 304 TyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIleIleProLeuArgTyrArg 323
 DB 1053 TATGAACAAATCTCTGAATTTGAAAAACGAATGCTCGTATTTATTTCCCTTCGTATCGT 1112
 QY 324 SerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSerThrProGlu 343
 DB 1113 TCAACCAATGGGTACCAGATTCAAGACCAAGAACCAAGTCCACCAACCGACTCCAGAA 1172
 QY 344 ProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAsp 363
 DB 1173 CTTAGTCCAGTCCGCAACCTGACCAATCTCAACAGCTCCAAGCAATCCAATTGAT 1232
 QY 364 GluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsn 383
 DB 1233 GAGAAATTTGGTCAAGAGAGCTGTTTCAAAAGTAGCGATGGTTATGCTTTTGGAGGAAAT 1292

; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 9
 ; LENGTH: 2451
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(2451)
 ; OTHER INFORMATION: n = a, c, t or g
 US-10-412-850-9

Alignment Scores:

Pred. No.: 6.83e-231 Length: 2451
 Score: 2658.50 Matches: 520
 Percent Similarity: 76.74% Conservative: 77
 Best Local Similarity: 66.84% Mismatches: 132
 Query Match: 66.15% Indels: 49
 DB: 16 Gaps: 8

US-09-765-272A-66 (1-763) x US-10-412-850-9 (1-2451)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
 Db 58 TGTCTTTACGAGTTGGGACTGTATCAAGCTAGAACGGTT---AAGGAAATAATCGTGT 114
 QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
 Db 115 TCTATATAGTGAACAAACAGCAGCAGAAACCGGAGAAATTTGACTCCCTGATGAGTT 174
 QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
 Db 175 AGCAAGCGTGAAGGAATCAATGCTGAGCAATCGTCATCAAGATAACAGCAAGCCTAT 234
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
 Db 235 GTCACTTTCATCGCGACCATATCATTTATTAACAATGTAAGTTCCCTATGACGCTATC 294
 QY 81 IleSerGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
 Db 295 ATCAGTGAAGAAATTAATCATGAAGATCCAAACTATAAGCTAAAGAGATGAGGATATTGT 354
 QY 101 AsnGluIleLysGlyLysValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
 Db 355 AATGAGTCAAGGTTGATATGTTATCAAGTAGATGGAATAACTATGTTTACCTTAAG 414
 QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
 Db 415 GATGCTGCCACGCGGATAACGTCGTAACAAAGAGGAAATCAATCGACMAAAACAAGAG 474
 QY 141 ArgSerHisAsnHisAsnSer-----ArgAlaAspAsnAlaValAlaAlaAlaArg 157
 Db 475 CATAGTCAACATCGTGAAGTGGAACTCCAGAAACGATGGTGTGCTTGCCTTGGCAGCT 534
 QY 158 AlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177
 Db 535 TCGCAGAGACGCTATPACACAGATGATGTTATCTTTAATGCTTCGATATCATAGAG 594
 QY 178 AspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsn 197
 Db 595 GATACTGGTATGTTATATCGTTCCTCATGAGATCATTAACCATATCATTCCTTAAGAT 654
 QY 198 GluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGly----- 213
 Db 655 GAGTTATCAGCTAGCGAGTTGGCTGCTGAGAAAGCCTTCCTATCTGCTCGAGGAATCTG 714
 QY 214 -----LysGlnGlySerArgProSerSerSerSerSerTyrAsn 226
 Db 715 TCAAAATCAAGAACCTATCGCCGCAAAAATAGCATACATCTCAAGAACCAACCTGGGTA 774
 QY 227 AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyr 246
 Db 775 CCTTCTGTAGCAATCCAGGAACCTACAAATACTATAACACAGCAACACAGCAACACTAAC 834

QY 247 HisGln---AsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysPro 265
 Db 835 AGTCAAGCAAGTCAAAAGTAATGACATTGATATCTCTTTGAAACAGCTCTACAAACTGCCT 894
 QY 266 LeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSer 285
 Db 895 TTGAGTCAACGACATGTAGAATCTGATGCGCTTGTCTTTGATCCACACAAATACAAAGT 954
 QY 286 ArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGlu 305
 Db 955 CGAACGCTAGAGGTGTGAGTGCACACGAGATCATTAACCATTCATCCCTTACTCT 1014
 QY 306 GlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsn 325
 Db 1015 CAATGTCTGAATTCGAAGAACGATCGCTCGTATTATTTCCCTTCCTTATCGTTCAAAC 1074
 QY 326 HisTrpValProAspSerArgProGlnProSerProGlnSerThrProGluProSer 345
 Db 1075 CATTTGGGTACCAAGATTCAAGGCGAGAACCAACCAAGTCCACACCGACTCCGGAACCTAGT 1134
 QY 346 ProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLys 365
 Db 1135 CCAGCCCGCAGACCTGCACCAATCTTAAA---ATAGACTCAAAAT-----TCTTCT 1182
 QY 366 LeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyVal 385
 Db 1193 TTGGTTAGTACGCTGTCGAGAAAGTTGGGGAAGGATATGTATTCGAGAAAGAGGCATC 1242
 QY 386 SerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys 405
 Db 1243 TCTCGTTATGCTTTGCGAAAGATTACCATCTGAAACTGTTAAAAATCTTGAAGCAAG 1302
 QY 406 LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSer 425
 Db 1303 TTATCAAAACAAGAGAGTGTTCACACACTTTAACTGCTAAAAAGAAATGTTGCTCCT 1362
 QY 426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445
 Db 1363 CGTGACCAAGAAATTTATGATAAGCATATATCTGTTAACTGAGGCTCATAAAGCCTTG 1422
 QY 446 LeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuArgLeu 465
 Db 1423 TTTGNAATAAGGTCGTAAATTCGATTTCCAGCCTTAGACAAATTTATTAGAAGCCTTG 1482
 QY 466 LysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaPro 485
 Db 1483 AATGATGAATCGCACTAATAAAGAAAAATTTGGTAGATGATTTATGGCAATTCCTAGCACCA 1542
 QY 486 IleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGlu 505
 Db 1543 ATTACCATCCAGAGCACTTGGCAAAACCAATCTCAAAATTCAGTATATCTGAAGCGAA 1602
 QY 506 IleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspPro 525
 Db 1603 GTTCGTTATGCTCAATTTAGCTGATAAGTATACACGCTCAGATGTTTACATTTTGTATGAA 1662
 QY 526 ArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHis 545
 Db 1663 CATGATATATCATGATGAAGGAGATGCATATGTAAACGCTCATATGGGCCCATAGTCAC 1722
 QY 546 TrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLys 565
 Db 1723 TGGATTGGAAAAAGATAGCCTTTCTGATAAGAAAAAGTTGCAGCTCAAGCCTATATCTAAA 1782
 QY 566 GluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLys 585
 Db 1783 GAAAAAGGTATCTTACTCTCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGAT 1842
 QY 586 GlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMet 605
 Db 1843 AGTGACGACGACTATTTTACAATCGTGTGAAGGGGAAAAACGAATTCCTCACTCGTTCGACTT 1902

366 LeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyVal 385
1183 TTGGTTAGTCAGCTGGTACGAAAGTTGGGAGAGATGATGATTCGAAGAAAGGCGATC 1242
386 SerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys 405
1243 TCTCGTTATGTCCTTGGCAAGATTACCATCTGAAACTGTTAAATAATCTTGAAGCAG 1302
406 LeuAlaLysGlnGluSerLysHisLysLeuGlyAlaLysLysThrAspLeuProSer 425
1303 TTATCAAAACAGAGAGTGTTCACACACTTTAACTGCTAAAGAAAGAAATGTTCTCT 1362
426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445
1363 CGTGACCAAGATTTATGATAGACATATATCTGTTAACTGAGGCTCATAAAGCCTTG 1422
446 LeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuGluArgLeu 465
1423 TTTGNAAAATAAGGTCGTATCTGATTCTCAAGCCTTAGACAAATTTATAGAGCCTTG 1482
466 LysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaPro 485
1483 AATGATGAATCCACTAATAAAGAAATTTGGTAGATGATTATTTGGCATTCCTAGACCA 1542
486 IleArgHisProGluAlaGluLysProAsnAlaGlnIleThrTyrThrAspAspGlu 505
1543 ATTACCATTCAGAGGACTTGGCAACAATTCCTCAATTCAGTATCTGAGACGAA 1602
506 IleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspPro 525
1603 GTTCGTATTGCTCAATTAGTAGTAAGTATACACGTCAGATGGTTACATTTTGTATGAA 1662
526 ArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHis 545
1663 CATGATATAATCAGTGTAGAGAGATGCATATGTAAACGCTCATATGGCCATAGTCTAC 1722
546 TrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGluAlaTyrAlaLys 565
1723 TGGATTGGAAGAAGTAGCCTTTCGTATAGGAAAGTTGCGAGCTCAAGCCCTATACTAAA 1782
566 GluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLys 585
1783 GAAAAAGGTATCCTACCTCCATCTCCAGAGCAGATGTTAAAGCAATCCAACTGGAGAT 1842
586 GlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMet 605
1843 AGTCAGCAGCATTTTACAATCGTGTGAAGGGGAAACGAATTCCTGCTGCTGACTT 1902
606 ProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyr 625
1903 CCATATATGTTGAGCATACAGTTGAGTTAAACCGTAATTAATTAATTTCTCTATAAG 1962
626 AspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLys 645
1963 GATCATTAACATAATTAATTAATTTGTTGTTGATGATCACACATACAAAGCTCCAAAT 2022
646 GlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665
2023 GGCTATACCTTGGAGATTTGTTTGGCAGATTAAGTACTAGTAGAACACCTCTGACGAA 2082
666 ArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn 685
2083 CGTCCACATCTTAATGATGATGGGCAATGCCAGTGCAGCATGTGTAGGCAAGAAAGAC 2142
686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
2143 CACAGTGAAGATCCAAATGAAGAACTTCAAGCGGATGAAGAG----- 2184
706 ProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725
2185 CCAGTAGAGGAACA----- 2199

726 ProThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGlu 745
2200 -----CCTGCTGAGCCAGAGTCCCTCAAGTAGAGACTGAA 2235
746 LysValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763
2236 AAAGTAGAAGCCCACTCAAGAAAGCAGAGTTTCTTCGAAAGTAGCGAT 2289
RESULT 16
US-09-765-272-55
; Sequence 55 Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-765-272-55
Alignment Scores:
Pred. No.: 4,31e-230 Length: 2389
Score: 2649.50 Matches: 519
Percent Similarity: 76.71% Conservative: 77
Best Local Similarity: 66.80% Mismatches: 132
Query Match: 65.92% Indels: 49
DB: 9 Gaps: 8
US-09-765-272A-66 (1-763) x US-09-765-272-55 (1-2389)

QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21
Db 2 TCTTACAGTTGGAGCTGTTATCAAGCTAGACGGTT---AAGGAAATATATCGTTTCC 58
QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSer 41
Db 59 TATATAGTGAAGAAACACGACGCAAAACCGAGAAATTTGACTCTCTGATGAGTTAGC 118
QY 42 LysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyrVal 61
Db 119 AAGCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAAGATAACAGACCAAGGCTATGTC 178

62 ThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIlelle 81
179 ACTTCACATGCGGACCATCATCATATTATTAACAATGGTAAGGTTCTCTTATGACGTATCATC 238
82 SerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsn 101
239 AGTGAAGAATTAATCATGAAGAATCCAAATCAATTAAGCTAAAGATGAGGATATGTGAAT 298
102 GluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAsp 121
299 GAGGTCAGGCTGGATATGTTATCAAGGTAGATGGAATAATCACTATGTTTACCTTAAGGAT 358
122 AlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGluArg 141
359 GCTGCCACCGCGATAACTCGCTACAAAGAGGAATCAATCGACAAAACCAAGAGCAT 418
142 SerHisAsnHisAsnSer-----ArgAlaAspAsnAlaValAlaAlaAlaArgAla 158
419 AGTCAACATCGTGAAGGTGGAACCTCCAGAGAAACGATGGTCTGTTCCTTGGCAGCGTTCG 478
159 GlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAsp 178
479 CAAGAGCGCTATACATACATGATGATGTTATGTTTAACTTCTGATATCATAGAGAT 538
179 ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGlu 198
539 ACTGGTGATGCTTATATGCTCTCATGAGATCAITACCATTACATTCCTTAAGATGAG 598
199 LeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGly----- 213
599 TTATCAGCTAGCGAGTGGCTGCTGAGAGAGCTTCTCTATCTGTCGAGGAAATCTGTCA 658
214 -----LysGlnGlySerArgProSerSerSerSerSerSerTyrAsnAla 227
659 AATTCAGAAACCTATGCGCAGAAAATAGCGATACATTTCAAGAACAACTGGGTACCT 718
228 AsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrHis 247
719 TCTGTAAGCAATCCAGGAATCAATACTAAACACAGCAACACACACTTAACAGT 778
248 Gln---AsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu 266
779 CAAGCAAGTCAAAGTAATGACATTTGATGCTCTTGAACACAGCTCTACAAACTGCCCTTG 838
267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286
839 AGTCAACGACATGATAGATCTGATGCGCTTGTCTTGTATCCCTTCTGTTATCGTCAAA 898
287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
899 ACAGCTAGAGGTGTCAGTGCACACGAGATCATTTACCCTTCATCCCTTTACTCTCAA 958
307 MetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHis 326
959 ATGCTCTGAATGGAAGACGAATCGCTCGTATATTCCCTTCTGTTATCGTTCAAACCAT 1018
327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346
1019 TGGGTACCATGATTTCAAGGCGAGAACCAACCAAGTCCCAACCGCATCCGGAACCTAGTCCA 1078
347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
1079 GCGCCGCAACCTGCACCAATCTTAA-----ATGACTCAAAAT-----TCTTCTTTG 1126
367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
1127 GTTAGTCAGCTGTGCGAAAAGTTGGGAAGATATGTTATCGAAGAAAGGCGATCTCT 1186
387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
1187 CGTTATGCTTTTTCGGAAGAGATTTACCATCTGAAATCTGTTAAATCTTGAAGCAAGTTA 1246
407 AlaLysGlnGluSerLeuSerHisLysLysLeuGlyAlaLysLysThrAspLeuProSerSer 426

1247 TCAAAACAAGAGAGTGTTCACACTTTAACTGTAATAAAGAAATGTTCTCTCTCGT 1306
427 AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeu 446
1307 GACCAAGATTTTATGATAAAGCATATATCTGTTAACTAGGCTCATAAAGCCTTGTTT 1366
447 AspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuGluArgLeuLys 466
1367 GNAATAAGGCTGCTAAATCTCTGATTTCCAGGCTTTAGACAAATTTATTAGACGCTTGAAT 1426
467 AspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIle 486
1427 GATGAATCGCATTAATAAAGAAATTTGATAGATGATTTATTGTCATTCCTAGCACAAT 1486
487 ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506
1487 ACCATCCAGAGCGACTTGGCAACAAATTTCTCAATTTAGTATGATCTAGACGCAAGT 1546
507 GlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArg 526
1547 CGTATTGCTCAATTAGCTGATAGTATATACAGCTCAGATGGTTACATTTTGTGTAACAT 1606
527 AspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTrp 546
1607 GATATAATCAGTATGAAGGAGATGATATGATACGCTCATATGGCCCATAGTCACTGG 1666
547 IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGlu 566
1667 ATTCGAAAAGATAGCTTTCTGATAAGAAAAGTTGCAGCTCAAGCTTATTAAGAA 1726
567 LysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGly 586
1727 AAAGTATCTTACCTTCCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGTAGT 1786
587 AlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMetPro 606
1787 GCACGAGCTATTTACATCGTGTGAAGGGGAAAACGAATTCCTCTGTCGACTTCCA 1846
607 TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAsp 626
1847 TATATGCTGAGCATACAGTTGAGCTTAAAAACCGTAATTTGATTTATTTCTCTCATAGGAT 1906
627 HisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLysGly 646
1907 CATTAACATAATATTTAAATTTGCTTGTGTTGATGATCACACATAAAGCTCCAAATGGC 1966
647 TyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg 666
1967 TATACCTTGGAGATTTGTTTGGACGATTAAGTACTAGTAGAAGACACCTGACGACGT 2026
667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGly 686
2027 CCACATTTCTAATGATGGATGGGCAATCCAGTAGCATGTGTAGGCAAGAAAGACCCAC 2086
687 GlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysPro 706
2087 AGTGAAGATCCAAATAAGAACTTCAAGCGGATGAAGAG-----CCA 2128
707 GluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLysPro 726
2129 GTAGAGGAAACA----- 2140
727 ThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLys 746
2141 -----CCTGCTGAGCGAGAGTCCCTCAAGTAGAGACTGAAAAA 2179
747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763
2180 GTAGAAGCCCAACTCAAAGAAGCAGAACTTTGCTTGGAAAGTAGTAACGAT 2230
RESULT 17
US-10-158-844-243

Sequence 243, Application US/10158844
Publication No. US20040029118A1
GENERAL INFORMATION:
APPLICANT: Kunsch et al.
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude Pentium 3
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10158,844
FILING DATE: 03-Jun-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/961,527
FILING DATE: 1997-10-30
APPLICATION NUMBER: US 60/029,960
FILING DATE: 1996-10-31
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB340P1D1
INFORMATION FOR SEQ ID NO: 243:
SEQUENCE CHARACTERISTICS:
LENGTH: 2359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 243:
US-10-158-844-243

Alignment Scores:
Pred. No.: 1.16e-209 Length: 2359
Score: 2424.00 Matches: 465
Percent Similarity: 97.50% Conservative: 3
Best Local Similarity: 96.88% Mismatches: 6
Query Match: 60.31% Indels: 6
DB: 13 Gaps: 1

US-09-765-272a-66 (1-763) x US-10-158-844-243 (1-2359)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValValLysGluSerAsnArgVal 20
DB 937 TGITTCCTATGAGCTTGACGCTTACCAAGCTGTCAGGATAGAAAGATCTTAATCGAGTT 996

QY 21 SerTyrTleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
DB 997 GCTTATATAGATGCTGATCATCGCTGGTCAAAGGCGAGAAACTTCACACCAGATGAAGTC 1056

QY 41 SerLysArgGluGlyLysAsnAlaGluGln***ValLleLysLleThrAspGlnGlyTyr 60
DB 1057 AGTAAGAGGGGGGGGATCAACGCCGACAAATTTGTTATCAAGATTACGGATCAAGGTTAT 1116

QY 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaLle 80
DB 1117 GTGACCTCTCATGAGACCATATATCAATACATAATGCAAGGTTCCCTATGATGCCATC 1176

QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspileVal 100
DB 1177 ATCACTGAGAGCTCCCATCAAGATCCGATATATCAATGAGGATTCAGAGATTGTC 1236

QY 101 AsnGluLysGlyGlyTyrValLleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
DB 1237 AATGAAATCAAGGGTGGTTATGTCATTAAAGTAAACGGTAAATATATATGTTTACCTTAAG 1296

QY 121 AspAlaAlaHisAlaAspAsnLleArgThrLysGluGluLysArgGlnLysGlnGlu 140
DB 1297 GATGACGCTCATCGGATATATTCGGACAAAGAGAGATTAAACCTCAGAACAGGAA 1356

QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
DB 1357 CGCAGTCATAATCATAACTCAAGAGCAGATAATGCTGTTGCTGCAGCCAGAGCCAAAGGA 1416

QY 161 ArgTyrThrThrAspAspGlyTyrLlePheAsnAlaSerAspileLleGluAspThrGly 180
DB 1417 CGTTATACCAACGGATGATGGGTATATCTTCAATGCATCTCATATCATTTAGAGCACCGGT 1476

QY 181 AspAlaTyrLleValProHisGlyAspHisTyrHisTyrLysProLysAsnGluLeuSer 200
DB 1477 GATGCTTATATCGTTCCTCAGCGCGACCATTAACATTACATTCTTAAGATGAGTTATCA 1536

QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgProSer 220
DB 1537 GCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGAAGCAGGATCTCGCTCTTCT 1596

QY 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
DB 1597 TCAAGTTCCTAGTTATTAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTG 1656

QY 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnLleSerSerLeuLeuArgGlu 260
DB 1657 ACTGTCACCTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAA 1716

QY 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuLlePheAspPro 280
DB 1717 TTGTATGCTAAACCCCTTATCAGAAACCCATGTGGAATCTGATGGCCTTATTTTCGACCCA 1776

QY 281 AlaGlnLleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
DB 1777 GCGCAAAATCACAAAGTCGAACCGCCAGAGGTGAGTGTCCCTCATGTGTAACTTACCATCA 1836

QY 301 PheLleProTyrGluGlnMetSerGluLeuGluLysArgLleAlaArgLleLleProLeu 320
DB 1837 TTTATCCCTTATGAACAAATGCTGAATTTGGAATAACGAATTTGCTCGTATATTTCCCTT 1896

QY 321 ArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSer 340
DB 1897 CGTTATCGTTCAAACCCATTTGGGTACAGATTCAAGACCAGAACCAAGTCCACACCG 1956

QY 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
DB 1957 ACTCCAGAAACCTAGTCCAAAGT-----CCGCAACCAAGCTCCCAAGCAAT 1998

QY 361 ProLleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
DB 1999 CCAATTTGATGAGAAATTTGTCAAAGAGAGCTGTTGCAAAAGTAGGCGATGTTATGCTTT 2058

QY 381 GluGluAsnGlyValSerArgTyrLleProAlaLysAspLeuSerAlaGluThrAlaAla 400
DB 2059 GAGGAGAAATGAGTTTCTCGTTATATCCAGCCAGAGATCTTTTTCAGCAGAAACAGCAGCA 2118

QY 401 GlyLleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLysGlyAlaLysLys 420
DB 2119 GGCATTGTATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAACTTAAGAAA 2178

QY 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
DB 2179 ACTGACCTCCCATCTAGTGATCGAGAAATTTTACATAAGGCTTATGACTTACTAGCAGA 2238

QY 441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
DB 2239 ATTCACCAAGATTTACTTCGATAAAGGTCGACAAAGTTGATTTTGGGCTTTTGGATAAC 2298

QY 461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspLleLeu 480
DB 2299 CTGTTTGGAAACGATCAAGGATGCTCTCAAGTGATAAAGTCAAGTTAGTGGGAAGATATCTT 2358

RESULT 18

JS-10-158-844-355

Sequence 355, Application US/10158844

Publication No. US20040029118A1

GENERAL INFORMATION:

APPLICANT: Kunsch et al.

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R

COMPUTER: Dell Latitude Pentium 3

OPERATING SYSTEM: Windows 98

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/158,844

FILING DATE: 03-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/961,527

FILING DATE: 1997-10-30

APPLICATION NUMBER: US 60/029,960

FILING DATE: 1996-10-31

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB340PID1

SEQUENCE CHARACTERISTICS:

LENGTH: 973 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 355:

JS-10-158-844-355

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

JB:

Length: 973

Matches: 266

Conservative: 2

Mismatch: 5

Indels: 1

Gaps: 0

2.73e-115

1377.00

98.17%

97.44%

34.26%

13

JS-09-765-272A-66 (1-763) x US-10-158-844-355 (1-973)

1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyValLysLysGluSerAsnArgVal 20

155 TGTTCTATGACCTGGAGCTTACCAAGCTGGTCAGATAAGAAAGCTATATCGATT 214

21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40

215 GCTTATATAGATGGTATCAGCTGGTCAAAAGGCAGAAAACCTTGACACAGATGAATC 274

41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60

275 AGTAAGAGGGAGGGGATCAACCGCGAACCAAAATGTTATCAAGATTACGGATCAAGGTTAT 334

61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80

335 GTGACCTCTCATGGAGACCATTAATCAATTAATGGCAAGGTTCTTATGATGCATC 394

81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysLysAspSerAspIleVal 100

395 ATCAGTGAAGAGCTCTCATGAAGATCCGAATATCAGTTGAAGATTACAGACATTGTC 454

101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120

Db 455 AATGAATCAAGGGTGGTTATGTCAATTAAGCTAAACCGTAATACTATGTTTACCTTAAG 514

QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140

Db 515 GATGAGCTCATGGGATATATTCGACAAAAGAGAGATTAACGTCAGAGCAGGAA 574

QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAraGalaGlnGly 160

Db 575 CGCAGTCATATCATTAATCAAGAGCAGATAATGCTGTGTCAGCCAGAGCCCAAGCA 634

QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180

Db 635 CGTTATACAACGGATGATGGTATATCTTCAATGCATCTGATATCATTCAGGACACGGGT 694

QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200

Db 695 GATGCTTATATCGTTCCTCAGCGGACCATTAACCATTAATCAATTCCTAAGAATGAGTTATCA 754

QY 201 AlaserGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgProSer 220

Db 755 GCTAGCGAGTTAGCTGCTGCAGAACCTTATGGATGGAGACGAGGATCTCGTCTTCT 814

QY 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240

Db 815 TCAAGTTCTAGTTATATGCAATCCAGCTCAACCAAGATTGTCAAGAACCAATCTG 874

QY 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260

Db 875 ACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAA 934

QY 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSer 273

Db 935 TTGTATGCTAA-CCCTTATCAGAACGCCATGTGGATCT 972

RESULT 19

US-09-884-465A-1

; Sequence 1, Application US/09884465A

; Publication No. US2003007293A1

; GENERAL INFORMATION:

; APPLICANT: Shire Biochem, Inc.

; APPLICANT: Hamel, Joseph

; APPLICANT: Brodeur, Bernard

; APPLICANT: Martin, Denis

; APPLICANT: Charland, Nathalie

; APPLICANT: Ouellet, Catherine

; TITLE OF INVENTION: Streptococcus Antigens

; FILE REFERENCE: 055190-0044

; CURRENT APPLICATION NUMBER: US/09/884,465A

; CURRENT FILING DATE: 2001-06-20

; PRIOR APPLICATION NUMBER: 60/212,683

; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 384

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 3120

; TYPE: DNA

; ORGANISM: Streptococcus pneumoniae

US-09-884-465A-1

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length: 3120

Matches: 296

Conservative: 101

Mismatch: 185

Indels: 149

Gaps: 24

US-09-765-272A-66 (1-763) x US-09-884-465A-1 (1-3120)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyValLysLysGluSerAsnArgVal 20

Db 61 TGTCCTATGACCTAAACCGATCGTTTCG---CAGGAAATAGGACATATATCGTCTC 117

QY	21	SerTyrIleAspGlyAspGlnAlaGlnLysAlaGluAsnLeuThrProAspGluVal	40
Db	118	TCVTATGTGATGCGACCAAGTCAAGTCAGAAAAAGTGAACACTTGCACACAGACCAAGT	177
QY	41	SerLysArgGluGlyIleAsnAlaGluGln***ValIleIysIleThrAspGlnGlyTyr	60
Db	178	AGCCAGAGAAGAGAAATTCAGCGCTGAGCAAAATGTAATCAAAATTACAGATCAGGGTAT	237
QY	61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
Db	238	GTAACGTCACACGGTGACCACTATCATTAATGGAAGAGTTCCTTATGATGCCCTC	297
QY	81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnIleLysAspSerAspIleVal	100
Db	298	TTTAGTGAAGAACTCTTGATGAGAGATCCAAACTATCAACTTAAGAAGCGCTGATATTGC	357
QY	101	AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120
Db	358	AATGAAGTCAAGGGTGTTATCATCTCAAGTCGATGGAATAATATATGCTACTCTGAAA	417
QY	121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
Db	418	GATGCAGCTCATGCTGATATGTTTCGAACATAAGATCAATCAATCGTCAAAAACAAGAA	477
QY	141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly	160
Db	478	CATGTCAAAAGAT---AATGAGAAGGTTAACTCTAAATGTTGCTGTAGCAAGGCTCTCAGGA	534
QY	161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
Db	535	CGATATACGCAAAATGATGGTTATGCTTTATUCCAGCTGATATATCGAAGATACGGGT	594
QY	181	AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer	200
Db	595	AATGCTTAATATCGTTCTCATGGAGGTCACCTATCACTACATTCCTCAAAAGCGATTATCT	654
QY	201	AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg	218
Db	655	GCTAGTGAATTACGACGAGCTAAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAG	714
QY	219	ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis	238
Db	715	TTAAGCTATTCTTCACAGCTAGTGACAAATAACACGCAG-----TCGTAGCAAAA	765
QY	239	AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu	258
Db	766	GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG	816
QY	259	ArgGluLeuTyrAlaLysProLeuSerSerGluArgHisValGluSerAspGlyLeuIlePhe	278
Db	817	AAGGAATCTTCATGATTCACTACGCGCCACAGTTTACAGTGAATCAGATGGCTGGTCTTT	876
QY	279	AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis	298
Db	877	GACCTGCTAAGATTATCACTCGTACACCAAAATGGAGTTGCGATTCGCGATGGGACCAT	936
QY	299	TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle	318
Db	937	TACCACCTTTATTCCTTACAGACAGCTTCTGCTTTAGAGAAAAAAGATTGCCAGAAATGGTG	996
QY	319	ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro	338
Db	997	CCATCAGTGGAACTGGTTCTACAGTTTCTACAAATGCAAAATCCAAAT-----1044	
QY	339	GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro	358
Db	1045	-----GAAGTAGTGTCTAGCTTAGCGACTCTTTCA	1074
QY	359	SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr	378
Db	1075	AGCAATCTCTTCTTTAAACGACGAATGAGAG---CTCTCTCTCAGCATCTGATGGTTAT	1131

RESULT 20

IS-09-884-465A-2

US-03-004-103A-2
: Sequence 2 Application US/09884465A

Sequence 2, Application US/0988871
Publication NO. US20030077382A1

; Publication No. US20
GENERAL INFORMATION

GENERAL INFORMATION:

; APPLICANT: Shire Biochem, Inc.

APPLICANT: Hamel, Josee

APPLICANT: Brodeur, Bernard

APPLICANT: Martin, Denis

APPLICANT: CHARL and NATHALIE

APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine

APPLICANT: Ouellet, Catherine

; TITLE OF INVENTION: Streptococcus Antigens

FILE REFERENCE: 055190-0044

; CURRENT APPLICATION NUMBER: US/09/88

CURRENT FILING DATE: 2001-06-20

CURRENT FILING DATE: 2001
PRIOR APPLICATION NUMBER:

; PRIOR APPLICATION NUMBER: 80/2
 : PRIOR FILING DATE: 2000-06-20

; PRIOR FILING DATE: 2000-06-20
 NUMBER OF SEQ ID NOS: 384

; NUMBER OF SEQ ID NOS: 384

; SOFTWARE: I

; SEQ ID NO 2

; LENGTH: 5048

Alignment Scores:

ATTACHMENT SCORES:
Pred No : 7 438-101

red. No.: 7.43e-101
1000 50
1000 50

Score: 1228.50

Percent Similarity: 54.31%

Best Local Similarity: 40.49%

Query Match: 30.57%

Length: 5048

Length:	5048
Matches:	200

Matches:	296
Games:	101

Conservative: 101

Mismatches: 185

Indels: 149

TS-09-765-272A-66 (1-763) x US-09-884-465A-2 (1-5048)

1	CysSerTyrGluLeuGlyArgHisglnAlaGlyGlnValbysLysGluSerAsnArgVal	20
1837	TGTGCTATGCATTAACCCAGCATCGTTCG--CAGGAAATAAGGACAATAATCGTGTC	1893
21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
1894	TCTTTATGTGGATGGCAGCAGTCACAGTCAGAAAGTGAAACCTTGACACAGACCAGGTT	1953
41	SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr	60
1954	AGCCAGAAAGAGGATTCAGGCTGAGCAATTTGTAATCAAAATTACAGATCAGGCGTAT	2013
61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
2014	GTAAGTCACAGCGTGACCACTATCATTAATATAATGGGAAGTTCCTTATGATGCCCTC	2073
81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal	100
2074	TTTATGTGAGAACTCTTGATGAAGAGTCCAAACTATCACTTAAAGACGCTGATATTGTC	2133
101	AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120
2134	AATGAAGTCAAGGCTGTTATATCATCAAGTCGATGAGAAATAATTATGCTACCTGAAA	2193
121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
2194	GATCGAGCTCATGCTGATATGTTCCAACTTAAAGATGAAATCAATCGTCAAAACACAGAA	2253
141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly	160
2254	CATGTCAAAGAT--AATGAGAAGGTTAACTCAATGTTCTGTAGCAAGGCTCTCAGGGA	2310
161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
2311	CGNTATACGCAAAATGATGGTTATGCTTTAAATCCAGCTGATATTTCAGAGATACGGGT	2370
181	AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer	200
2371	AATGCTTATATCGTCTCATGAGGTCATCTATCATATCCAAACGCGTTATCT	2430

```
QY 559 AlaAlaGlnAlaTyrAlaLysGlu-----LysGlyLeuThrProSer 573
Db 3214 GCTCGCAAAACATCTTACAGAGAGTTAAACTAGCTATAATGATGATCTTTGTCA 3273
QY 574 ThrAspHisGlnAsp---SerGlyAsnThr---GluAlaLysGlyAlaGluAlaIleTyr 591
Db 3274 TCTCATGACAGAGATTATCCAGGTATATCCCAAGAAATGAAGATTAGATAAA----- 3327
QY 592 AsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGln--- 610
Db 3328 -----AAAATCGAAGAAAAAATTCGTGCATTATGCAACAATATGGTGTCAACGT 3378
QY 611 -----TyrThrValGluValLysAsnGlySerLeuIlelleProHisTyrAspHisTyr 628
Db 3379 GAAAGTATTGCTGATTAATAAGAAAAAATCGGATTATTTATCCGATGGAGATCACCAT 3438
QY 629 His-----AsnIleLysPhe 633
Db 3439 CATGCAGATCCGATGATGAACATTAACCGGTTGGAATGTCATCTCACAGTAACTAT 3498
QY 634 GluTrpPhe-----AspGluGlyLeu-----TyrGluAlaProLysGlyTyrThrLeu 649
Db 3499 GAACGTGTTTAAACCCGAAGAGAGGAGTTGCTAAAGAAAGAGGGAATAAAGTTTATCTGGA 3558
QY 650 GluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSer 669
Db 3559 GAAGAAATTAACGAATGTTGTTAAATTTGTTA-----AAA 3591
QY 670 AspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAsp 689
Db 3592 AATAGTACGTTTAAATCAAAACCTTACTCA-----GCCAATGTCATAAAACGC 3642
QY 690 ThrAsnGlnThrGluLysProSerGluGluLys 700
Db 3643 GTTTCCTTTTACTTTTCGCGCTCAATTTGAGAAA 3675

RESULT 21
US-10-158-844-192/c
; Sequence 192, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340PID1
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6867 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-10-158-844-192

Alignment Scores:
Pred. No.: 2,57e-98 Length: 6867
Score: 1202.50 Matches: 294
Percent Similarity: 54.04% Conservatives: 101
Best Local Similarity: 40.22% Mismatches: 187
Query Match: 29.92% Indels: 150
DB: 13 Gaps: 23

US-09-765-272a-66 (1-763) x US-10-158-844-192 (1-6867)

QY 1 CysSerThrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 6725 TGTGCTATGCTACTAAACCCAGCATCGTTCG---CAGGAAATATAGGCAATATCGTGTC 6669
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 6668 TCTTATGTGATGGCAGCCAGTCAGTCAGAAAAGTGAAAACCTTGACACCCAGACCGGTT 6609
QY 41 SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr 60
Db 6608 AGCCAGAAAGAGGAATTCAGGCTGAGCAAAATGTATCAAAATTACAGATCAGGGCTAT 6549
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 6548 GTAACGTCACACGTCACCACTATCATCTATATAATGGGAAAGTTCTTATGATGCCCTC 6489
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 6488 TTTAGTGAAGAACTCTTTGATGAAGGATCCAACTATCAACTTAAAGACGCTGATATGTC 6429
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 6428 AATGAAGTCAAGGCTGTTATATCATCATGCTCGATGGAATAATATATGTCTACCTGAAA 6369
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 6368 GATGCACTCATGCTCATATGATGTCGAACATAAGATGAATCAATCTGTCAAAAACAGAA 6309
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
Db 6308 CATGTCAAAGAT---AATGAGAAGTTAACTCTAATGTTGCTGTAGCAGGTCTCAGGGA 6252
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 6251 CGATATACGACAAATGATGTTATGCTTTTAAATCCAGCTGATATATCGAAGATACGGGT 6192
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 6191 AATGCTTATATCGTCTCTCATGAGGTCACTATCATCTACATTCCTCCAAAGCGGATTTATCT 6132
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLys-----GlnGlySerArg 218
Db 6131 GCTAGTGAATTAGCAGCAGCTAAAGCACATCTCGCTGGAAAAAATATGCAACCGAGTCAG 6072
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 6071 TTAAGCTATTCTTCAACAGCTAGTAGCAATAACACGCAA-----TCTGTAGCAAAA 6021
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
Db 6020 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGCTTTTG 5970
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 5969 AAGGAACCTTATGATTCACCTAGCGCCACAGTGTACAGTGAATCAGATGCGCTGCTTT 5910
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 5910
```

5909 GACCTGCTAAGATTATCATGCTGTACACCAAAATGGAGTTCCGATTCGGATCCGATGGACCAT 5850
299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle 318
5849 TACCACCTTTATTCCTTACAGCAGCTTCTGCTGTAGAGAAAAAGATTGCCAGAATGGTG 5790
319 ProLeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerPro 338
5789 CTTATCAGTGGAACTGGTCTTACAGTTTTCACAAATGCAAAACCTAAT 5742
339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
5741 5742
359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
5711 AGCAATCCCTCTCTTTAAGCAGACAGTAAGAG--CTCTCTTCAGCATCTGATGGTTAT 5655
379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
5654 ATTTTAAAT 5622
399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAla 418
5621 GCTACA-----CCTATATTGTAAGCATGGTGATCATTTCCATTAC----- 5580
419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
5579 -----ATTCCAAATCAAT----- 5565
439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
5564 ---CAAATTGGGCAACCGACTCTTCCAAACAATAGTCTAGCAACCTCTCCATCTCTT 5508
459 AspAsnLeuLeuGluArgLeuLysAspVal**SerAspLysValLysLeuVal**Asp 478
5508 ----- 5508
479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
5507 -----CCAATC----- 5493
499 IleThrTyrThrAspAspGluLeGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
5492 ACTTCACATGAGAAACATGAA-----GAA 5469
519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
5468 GATGATACGGATTGATGCTAATCGTATATATCGTGAAGATGAATCAGGTTTGTCTCATG 5409
539 ProHisMetThrHisSerHisTyrPheLysLysAspSerLeuSerGluAlaGluArgAla 558
5408 AGTCACGGAGACCAATCATTTATTTCTTCAAGAGGACTTGACAGAGCAAAATTA-- 5350
559 AlaAlaGlnAlaTyrAlaLysGlu-----LysGlyLeuThrProProSer 573
5349 GSTGCGCAAAACATTAGAGGAAGTTAAACCTAGTCATAATGATTAGATTCTTTGTCA 5290
574 ThrAspHisGluAsp-----SerGlyAsnThrGluAlaLysGlyAlaGluAlaLeTyr 591
5289 TCTCATGACAGGATTATCCAAAGTAAATGCAAAAGAAATGAAGATTAGATAAA----- 5236
592 AsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGln--- 610
5235 -----AAATCGAAGAAAAATTTGCGCATTTATGAAACAATATGTTGTCGAACCGT 5185
611 -----TyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAspHisTyr 628
5184 GAAAGTATTGCTGTAATAAGAAAAAATGCGATTATTTATCCGATGAGATCACCAT 5125
629 His-----AsnIleLysPhe 633
5124 CATGCAGATCCGATTGATGAACATAAACCGTTGGTAATGGTCAATTCATCTCACAGTAACAT 5065

Qy 634 GluTrpPhe-----AspGluGlyLeu-----TyrGluAlaProLysGlyTyrThrLeu 649
Db 5064 GAACGTGTTTAAACCCGAGAGAGAGTGTCTAAAAAAGAGGAATTAAGTTTATCTGGA 5005
Qy 650 GluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSer 669
Db 5004 GAAGAATTAAACGAATGTTGTTAATTGTTA-----AAA 4972
Qy 670 AspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAsp 689
Db 4971 AATAGTACGTTTAATATCAAACTTTACTCTA-----GCCAATGGTCAAAACGC 4921
Qy 690 ThrAsnGlnThrGluLysProSerGluGluLys 700
Db 4920 GTTCTCTTATGTTTCCGCTGAATGGAGAAA 4888

RESULT 22

US-10-158-844-258
; Sequence 258, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340PLD1
; INFORMATION FOR SEQ ID NO: 258:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 258:
US-10-158-844-258

Alignment Scores:

Pred. No.:	6,36e-99	Length:	1684
Score:	1200.00	Matches:	228
Percent Similarity:	99.56%	Conservative:	0
Best Local Similarity:	99.56%	Mismatches:	1
Query Match:	29.86%	Indels:	1
DB:	13	Gaps:	0

US-09-765-272A-66 (1-763) x US-10-158-844-258 (1-1684)

Qy 535 AlaTyrValThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGlu 554
Db 3 GCCTATGTAATCCACATATGACCATGACCTAGCCACTGGATTAAAAAAGATAGTTTGTCTGAA 62

555	AlaGluArgAlaAlaAlaGlnAlaIleValAlaLysGluLysGlyLeuThrProSerThr	574
Db		
63	GCTGAGAGCGCGACC-CAGCGTTATGCTAAAGAGAAGGTTTGACCCCTCCTTCGACA	121
Db		
575	AspHisGlnAspSerGlyAsnThrGluAlaLysGlyValAlaGluAlaIleTyrAsnArgVal	594
Db		
122	GACCATCAGGATTCAGGAAATACTGAGGCMAAAGAGCGAAGACTATCTACAAACCGCGTG	181
Db		
595	LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu	614
Db		
182	AAAGCAGCTAAAGAGGTGCCACTTCATCGTAGCGCTTACAATCTTCAATATACTGTAGAA	241
Db		
615	ValLysAsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGlu	634
Db		
242	GTCAAAAACGGTAGTTTAAATCAATCACTCATTAAGACCAATTACCAATCAATCAAAATTGAG	301
Db		
635	TrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAla	654
Db		
302	TGGTTTGACGAAGGCGCTTTATGAGGCACCTAAGGGGTATACTCTTGAGGATCTTTGGCG	361
Db		
555	ThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGly	674
Db		
362	ACTGTCGAAGTACTATGTCGAACATCCAAACGAAACGTCGCGCATTCAGATAATGGTTTGGT	421
Db		
675	AsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGlu	694
Db		
422	AACGCTAGCGACCATGTTCAAAAGAAACAAAATGGTCAAGCTGATACCATAATCAAAACGGAA	481
Db		
695	LysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArgGluGlu	714
Db		
482	AAACCAAGCGGAGAAACCTCAGACAGAAAAACCTGAGGAGAAACCCCTCGAGAAAGAG	541
Db		
715	LysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSerPro	734
Db		
542	AAACCGCAAAGCGAGAAACCGAGGTCTCCAAAACCAACAGAGGAAACCGAGAAATCACC	601
Db		
735	GluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGluAla	754
Db		
602	GAGGAATCAGAGAACTCAGTTCAGTCAAGTCAAAAGGTTGAAGAAAAAATCTGAGAGAGCT	661
Db		
755	GluAspLeuLeuGlyLysIleGlnAsp	763
Db		
662	GAAGATTACTTGGAAAAATCCAGAT	688
Db		

RESULT 23

RESULT 23
US-09-769-787-246
; Sequence 246, Application US/09769787
; Publication No. US20030091577A1

```

1  / GENERAL INFORMATION:
2  / APPLICANT: Microbial Technics Limited
3  / APPLICANT: Gilbert, Christophe FG
4  / APPLICANT: Hansbro, Philip M
5  / TITLE OF INVENTION: Proteins
6  / FILE REFERENCE: PWC/P21129WO
7  / CURRENT APPLICATION NUMBER: US/09/769,787
8  / CURRENT FILING DATE: 2001-01-26
9  / PRIOR APPLICATION NUMBER: GB 9816337.1
10 / PRIOR FILING DATE: 1998-03-27
11 / PRIOR APPLICATION NUMBER: US 60/125164
12 / PRIOR FILING DATE: 1999-03-19
13 / NUMBER OF SEQ ID NOS: 388
14 / SOFTWARE: PatentIn Ver. 2.1
15 / SEQ ID NO 246
16 / LENGTH: 1455
17 / TYPE: DNA
18 / ORGANISM: Streptococcus pneumoniae
19 / US-09-769-787-246

```

Alignment Scores:	
Pred. No.:	4.6e-96
Score:	1167.50
Percent Similarity:	59.50%
Best Local Similarity:	45.52%
Length:	1455
Matches:	254
Conservative:	78
Mismatches:	123

	Query Match:	29.05%	Indels:	103
	DB:	10	Gaps:	14
<hr/>				
US-09-765-272A-66 (1-763) x US-09-769-787-246 (1-1455)				
QY	1	CysSerTyrGluLeuGlyVArHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal	20	
Dd	61	TGTGGCTATGCACATAAACCCAGCATCGTTTCG-----CAGGAATAAGCAATAATCGTGTC	117	
QY	21	SerTyriLeAspGlyVsPglNalacgLyGlnLysAlaGluAsnLeuThrProAspGluVal	40	
Dd	118	TCTTATGTGGATGGCAGCGCATCAAGTAAGAAGTGAAAATTGACAACCAGCACGGTT	177	
QY	41	SerIysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr	60	
Dd	178	AGCCAGAAAGAAGAAATTCAGGCTCAGCAAATTGTAATCAAAATTTACAGATCAGCGCTAT	237	
QY	61	ValThrSerHisGlyAspHisTyHisTyTYrAsnGlyLysValProTYrAspAlarile	80	
Dd	238	GTAACGTACACGGTAGCACTATCATTTACTATTAATGGGAAGTTCCTTATGATGCCCTC	297	
QY	81	IleSerGluGluLeuLeuMetLysAspProAsnTYrGlnLeuLysAspSeraSpileVal	100	
Dd	298	TTTTAGTGAAGAACTCTTCATGAAGATCCAACCTATCAACTTAAGACGCTGATATTGTC	357	
QY	101	AsnGluuleLysGlyTYrTYrValilleLysValasnGlyLystyTYrValTYrneuLys	120	
Dd	358	AATGAAGTCAAGGGTGGTTATCATCAAGGTCGATGGAAAAATATTATGTCTACTGAAA	417	
QY	121	AspAlaalHiSaLaAspAsenlleArgThrylsGluGlutlileyAsrgGlnlysGlnGlu	140	
Dd	418	GATCGACTCATGCTGATAATGTTTCGAACATAAAGATGAATCAATCGTCAAAAACAAGAA	477	
QY	141	ArgSerHisAsnHisAsnSerArgAlaaspAsnAlaValalaAlaAlaArgAlaglnGly	160	
Dd	478	CATGTCAAAGAT---AATGAGAAGTTAACTCTAATGTTCGTGTAGCAAGTCTCAGGGA	534	
QY	161	ArgTYrThrThrAspAspGlyTYrilPhaAsnAlaSerAspilelleGluAspThrgly	180	
Dd	535	CGATPATAGCAAAATGATGGTTATGTCTTTAACTCCAGCTGATATTATCGAAGATACGGGT	594	
QY	181	AspAlatvriLeValProHIsAglyAspHisHyHisTyriLeProLYsaSnGluLeuSer	200	
Dd	595	AATGCTTATATCGTTCTCTCATGGAGGTCACATCATCATCTCCAAAAGAGGATTTATCT	654	
QY	201	AlsSerGluLeuAlaAlaGluAlatyTrpAsnGLYLys-----GinglySerArg	218	
Dd	655	GCTAGTCAATTAGCAGCAGCTAAAGCACATCTGGCTGMAAAAATATGCAACCGAGTCAG	714	
QY	219	ProSerSerSerSerTYrAsnAlaAsnProAlaGlnProArgLeuSerGluasNHIS	238	
Dd	715	TTTAGCTATTCTTCAACAGCTAGTGACAATAACACGCAG-----TCTGTAGCAAAA	765	
QY	239	AsnLeuThrValThrProTHrTyHisGlnAsnPglngLyGluasnlleSerSerLeuLeu	258	
Dd	766	GGATCAACTAGCAGCCAGCA-----AATAAATCTGAANAATCTCCAGAGTCITTTG	816	
QY	259	ArgGluLeuTyralalySProLeuSerGluARhisValGluSerAspGlyLeuillePhe	278	
Dd	817	AAGGAACCTCATGATTCAACCTAGGCCCAACGTTACAGTGAATCAGATGCGCTGGTCTTT	876	
QY	279	AspProAlaclnlethrSerArgGthrAlaAargGlyValalavalProHisGlyAsnHis	298	
Dd	877	GACCTGCTAAGATTATCAGTCGTACCAAAATGGAGTTGCGATTCCGATGCGACCAT	936	
QY	299	TyrHisPheilleProTYrGluGlnMetSerGluLeuGlynusArgtileAlaArgilleile	318	
Dd	937	TACCATTATCTTACAGCAAGCTTTCTGCTTTAGAGAAAAGATGCCAAGATGGTG	996	
QY	319	ProLeuArgTYrArgSerAsnHisTrpValProAspSeraArgProGluGlnProSerPro	338	
Dd	997	CCATCAGTGGAACTGGTCTACAGTTCTTACAAAATGCAAAACCTAAT-----1044		

2Y 339 GlnSerThrProGluProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAAGTAGTGTCTAGCTAGGCGAGCTTTTCA 1074
2Y 359 SerAsnProIleAspGluLysLeuValLysGlnAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCTTCTTTTAAACGACAGTAAGAG---CTCTCTTCCAGCATCTGATGGTTAT 1131
2Y 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 ATTTTAAAT-----CCAAAGATATCGTTGAAGAACG 1164
2Y 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysHisLysLeuGlyAla 418
Db 1165 GCTACA-----GCTTATATTGTGAAGCATGGTGATCATTTCCATTAC----- 1206
2Y 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAATCAAAAT----- 1221
2Y 439 AlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAATTCGGCAACCGACTCTTCCAAACATAGTCTAGCAACACCTTCTCCATCTCTT 1278
2Y 459 AspAsnLeuLeuGluArgLysAspVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
2Y 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CCAATC-----AATCCAGGA 1293
2Y 499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
2Y 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GAUGGATACGGATTGATGCTAATCGTATTATCGCTGAAGATGAATCAGGTTTGTCTATG 1377
2Y 539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGlu 556
Db 1378 AGTCACGGAGACCAACATATTATTTCTTCAAGAAGGACTTGACAGAAGAGCAA 1431

RESULT 24
JS-09-769-744A-23
Sequence 23, Application US/09769744A
Publication No. US20030134407A1
GENERAL INFORMATION:
APPLICANT: Le Page, Richard WF
APPLICANT: Wells, Jeremy M
APPLICANT: Hanniffy, Sean B
APPLICANT: Hansbro, Philip M
TITLE OF INVENTION: Proteins
FILE REFERENCE: PWC/P21122WO
CURRENT APPLICATION NUMBER: US/09/769, 744A
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: PCT/GB99/02452
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: GB 9816336.3
PRIOR FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: US 60/125329
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 196
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 23
LENGTH: 1455
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
JS-09-769-744A-23

Alignment Scores: 4.6e-96 Length: 1455
Pred. No.:

Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.53% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 10 Gaps: 14
US-09-765-272A-66 (1-763) x US-09-769-744A-23 (1-1455)
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
Db 61 TGTGCTATGCTAAACCGCATCTGTCG---CAGAAATTAAGACATAATTCGTGTC 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGATGGTGGCAGCCAGTCAAGTCAGAAAGTGAACCTTGACACCAGACAGTT 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAGGAAATTCAGGCTGAGCAAAATGTAATCAAAATTAAGATCAGGCTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTACGTCACAGGTCAGCCACTATCATTAATATGGAAGTTCCTTATGATGCCCTC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysAspSerAspIleVal 100
Db 298 TTTAGTGAAGAACTCTGATGAAGGATCCAACTATCACTTAAGACGCTGATATGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 358 AATGAAGTCAAGGGTGGTTATATCATCAAGTCGATGGAATATATATGCTACTCTGAAA 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnGluGlu 140
Db 418 GATGCGCTCATGCTGATAATGTTTCAACTAAGATGAATCAATCGTCAAAACCAAGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaGlnGly 160
Db 478 CATGTCAAAGAT---AATGAGAGGTTAACTTAATGTTGCTGTAGCAGGCTCAGGGA 534
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACGACAAATGATGTTATGCTTTAATCCAGCTGATATATTCGAAGATACGGGT 594
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTTCTCTCATGGAGGTCATCATCACTACATTCCTCCAAAGCGATTTATCT 654
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLys---GlnGlySerArg 218
Db 655 GCTAGTGAATTAAGCAGCAGCTAAGCAGCATCTGGCTGGAAATAATATGCAACCGAGTCAG 714
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGTATTCTTCAACAGCTAGTGAACAATAACACGCA---TCTGTAGCAAAA 765
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnLysSerSerLeuLeu 258
Db 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAATAATCTCCAGAGTCTTTTG 816
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuPhe 278
Db 817 AAGAACTCTATGATTCACCTAGGCCCAAGCTTACAGTGAATCAGATGGCTGCTTTT 876
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCTGCTAAGATTATCAGTCGTACACCAATGAGGTTGCGATTCCGATCGGACCAT 936
QY 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle 318
Db 937 TACCACCTTATCTTACAGCAAGCTTCTGCTTCTAGAAAGAAAGATTGCCAGATGGTG 996
QY 319 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338

```

Db 997 CCTATCAGTGGAACTGGTTCTACAGTTTCTACAAATGCAAAACCTTAAT----- 1044
Qy 339 GlnSerThrProGluProSerProSerProGlnProAlaProGlnProAlaPro 358
Db 1045 -----GAGTAGTGTCTAGCTAGGAGTCTTTCA 1074
Qy 359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCCTTCTTTTAAACGACAAAGTAAGAG-CTCTCTTCCAGCATCTGATGGTTAT 1131
Qy 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 AATTTTAAT-----CCAAAAGATATCCTTGAAGAAGCG 1164
Qy 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysSerHisLysLeuGlyAla 418
Db 1165 GCTACA-----GCITATATTGACACATGGTGATCAATTCCTTAC----- 1206
Qy 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAAATCAAAT----- 1221
Qy 439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAAATGGCAACCGACTCTTCCAAACAATAGTGTAGCAACACTTCTCCATCTCTT 1278
Qy 459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
Qy 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CCAATC-----AATCCAGGA 1293
Qy 499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
Qy 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GATGATACCGATTGTGCTAATCGTATTATCGCTGAAGATGATCAGGTTTGTTCATG 1377
Qy 539 ProHisMetThrHisSerHisTyrIleLysLysAspSerLysSerGluAlaGlu 556
Db 1378 AGTCACGAGAGACCAATCAATATTCTTCAAGAGGACTTGCACAGAGAGCAA 1431
RESULT 25
US-10-412-862-7
; Sequence 7, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-862-7
Alignment Scores:

```

```

Pred. No.: 4,6e-96 Length: 1455
Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.52% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 13 Gaps: 14

```

US-09-765-272A-66 (1-763) x US-10-412-862-7 (1-1455)

```

Qy 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
Db 61 TGTGCTTATCCATTAACAGCATCGTTTCG---CAGGAAATTAAGCAATAATATCGTGTCT 117
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGTGATGGCAGCCAGTCAAGTCAGAAAAGTGAAAACCTTGACACCAAGCAGGTT 177
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAGGAATTCAGGCTGAGCAAAATTGTAATCAAAATTACAGATCAGGGCTAT 237
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTAACGTCAACCGGTGACCACTATCATCTATAATGGGAAAGTTCTTATGATGCCCTC 297
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 TTTAGTGAAGAACTCTTGATGAAGGATCCAACTATCACTTAAAGCGGTGATTTGTC 357
Qy 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValLysLeuLys 120
Db 358 AATGAAGTCAAGGCTGGTTATATCATCATCAAGGTCGATGGAATAATATTATGCTCTCTGAAA 417
Qy 121 AspAlaHisAlaAspAsnIleArgThrLysGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGCGAGCTCATGCTGATATGTTCCGAACATAAGATGAAATCAATCTGCAAAAACAAGAA 477
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaAlaGlnGly 160
Db 478 CATGTCAAAGAT---AATGAAAGGTTAACTTAATGTTGCTGTAGCAAGGCTCAGGGA 534
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACACAAATGATGTTATGCTTTAATCCAGCTCATATTATCGAAGATACGGGT 594
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTTCTCATGAGGTCACTATCACTACATTCCTCCAAAAGCGATTTATCT 654
Qy 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLys-----GlnGlySerArg 218
Db 655 GCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGTGGAAAAAATAATGCAACCGAGTCAG 714
Qy 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGCTATTCTTCAACAGCTAGTGACATAAACAACGCAA-----TCTGTAGCAAAA 765
Qy 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
Db 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG 816
Qy 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 817 AAGGAACTCTATGATTCACCTAGCGCCCAACGGTTACAGTGAATCAGATGCCCTGGTCTTT 876
Qy 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCTGCTAAGTATATCATGTCACCAATGGAGTTGCGATTCCGATGCGGACCAT 936
Qy 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle 318
Db 937 TACCACCTTATTCCTTACAGCAAGCTTCTGCTTAGAAGAAAAAGATTGCGAATCGTG 996

```

```

319 ProLeuArgTyrArgSerAsnHisTipValProAspSerArgProGluInProSerPro 338
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
997 CCTATCAGTCGAAGTGGTTCTACAGTTCTTACAAATGCAGAACTTAAT----- 1044
339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1045 -----GAAGTAGTGTCTAGCTAGGAGCTTTTCA 1074
359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1075 AGCAATCTCTTCTTTTAAACACAGTAAGGAG--CTCTCTTCCAGCATCTGATGGTTAT 1131
379 valPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1132 ATTTTAAAT-----CCAAAGATATCTGTTGAAGAAACG 1164
399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAla 418
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1165 GCTACA-----GCTTATATTGTAAGACATGGTGATCATTTCCATTAC----- 1206
419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1207 -----ATTCAAAATCAAT----- 1221
439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1222 -----CAAATTTGGGCAACGCTCTTCCAAACATAGTCTAGCAACACCTTCTCCATCTCTT 1278
459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1278 ----- 1278
479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1279 -----CCAATC----- 1293
499 IleThrTyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1294 ACTTCACGAGCAACATGAA-----GAA 1317
519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1318 GATGATACGGAATTTGATGCTTAATCGTATATATCGCTGAAGATGAATCAGGTTTGTCTATG 1377
539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLysLeuSerGluAlaGlu 556
|||...:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 AGTCACGGAGACCAATCATATTATTTCTTCAGAAAGGACTTGACAGAGAGCAA 1431

```

RESULT 26

IS-10-412-850-7

Sequence 7, Application US/10412850

Publication No. US20040001836A1

GENERAL INFORMATION:

APPLICANT: Johnson, Leslie S.

APPLICANT: Adamou, John E.

TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus

TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural

TITLE OF INVENTION: Motifs

FILE REFERENCE: 469201-686

CURRENT APPLICATION NUMBER: US/10/412,850

PRIOR FILING DATE: 2003-04-14

PRIOR APPLICATION NUMBER: 09/468,656

PRIOR FILING DATE: 1999-12-21

PRIOR APPLICATION NUMBER: 60/113,048

PRIOR FILING DATE: 1998-12-21

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 3.0

SEQ ID NO 7

LENGTH: 1455

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

IS-10-412-850-7

Alignment Scores:

```

Pred. No.: 4,6e-96 Length: 1455
Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.52% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 16 Gaps: 14

```

US-09-765-272A-66 (1-763) x US-10-412-850-7 (1-1455)

```

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
DB 61 TGTGCTTATGCACTAAACACGATCGTTCCG---CAGGAAAAATTAAGACATAATAATCGGTGC 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyLysAlaGluAsnLeuThrProAspGluVal 40
DB 118 TCTTATGTGGATGGCAGCCGCTCAAGTCAGAAAGTGAAACCTTGACACGACCGAGGTT 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
DB 178 AGCCAGAAAGAGGAATTCAGGCTGAGCAAAATGTGAATCAAAATTACAGATCAGGGCTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
DB 238 GTAACGTCACACGGTGACCACTATCATTAATAATGGAAGTTCCTTATGATGCCCTC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
DB 298 TTTAGTGAAGAACTCTTGATGAAGGATCCAACTATCACTTAAAGACGCTGATATTGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
DB 358 AATGAAGTCAAGGGTGGTTATATCATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
DB 418 GATGCACTCATGCTGATAATGTTTCAACTAAAGATGAATCAATCGTCAAAAACAAGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
DB 478 CATGTCAAAAGAT---AATGAGAAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGGA 534
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
DB 535 CGATATACGACAAATGATGGTTATGCTTTTATCCAGCTGATATATATCGAAGATCGGTT 594
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
DB 595 AATGCTTTATATCGTTCTCTCATGGAGTCACTATCATCTACATTCCTCCAAAAGCGATTATCT 654
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg 218
DB 655 GCTAGTGAATTTAGCAGCAGCTAAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAG 714
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
DB 715 TTAAGCTATTCTTCACAGCTAGTAGACAAATAACGCGCAA-----TCTGTACAAAA 765
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
DB 766 GGATCAACTAGCAGCCAGCA-----AATAAATCTGAAATCTCCAGAGCTTTTGTG 816
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
DB 817 AAGGAATCTTATGATTCACCTAGCGCCCAAGTTACAGTGAATCAGATGGCGCTGCTCTT 876
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
DB 877 GACCTGCTAAGATTATTCAGTGTACACCAAAATGGAGTTGCGATTCGCGATGGGACCAT 936
QY 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle 318
DB 937 TACCACTTTATCTTACAGCAAGCTTTCTGCCTTAGAAGAAAAGATTGCCAGATGGTG 996

```



```
QY 319 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338
Db 997 CCTATCAGTGAACACTGGTCTACAGCTTTCTACAAATGCAAAACCTTAAT----- 1044
QY 339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAAAGTAGTGCTAGTCTAGGAGAGCTTTTCA 1074
QY 359 SerAsnProIleAspGluLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCTCTCTTTAAACGACAAAGTAGGAG---CTCTCTCAGCATCTCATGCTTAT 1131
QY 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 ATTTTAAAT-----CCAAAAGATATCGTTGAAGAAACG 1164
QY 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLysGlyAla 418
Db 1165 GCTACA-----GCTTATATTCTAAGACATGCTGATCATTTCCATTAC----- 1206
QY 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAATCAAT----- 1221
QY 439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAAATTTGGCAACCGACTCTTCCAAACATAGCTAGCAACACCTTCTCCATCTCTT 1278
QY 459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
QY 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CCAATC-----AATCCAGCA 1293
QY 499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
QY 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GATGGATACGGATTGTGCTGAATCGTATTATCGCTGAAGATGAATCGATTGTCATG 1377
QY 539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGlu 556
Db 1378 AGTCACGGAGACCAATCATATTCTTCAAGAAGGACTTGACAGAGAGCAA 1431
RESULT 27
US-10-387-783-7
; Sequence 7, Application US/10387783
; Publication No. US2004000531A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-387-783-7
```

Alignment Scores:

```
Pred. No.: 4.6e-96 Length: 1455
Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.52% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 16 Gaps: 14
```

US-09-765-272A-66 (1-763) x US-10-387-783-7 (1-1455)

```
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 61 TGTGCGCTATGCACTAAACCCAGCATCGTTCG---CAGGAAATAAGCAATAATTCGTGTC 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGTGGTGGCCGCCAGTCAAGTCAGAAAAGTGAAAACCTTGACACCCAGACAGGTT 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAAAGAAATTCAGGCTGAGCAAAATTGTAATCAAAATTACAGATCAGGCTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTAACGCTCACCGGTGACCACTATCATATAATGGAAGAGTTCTTATGATGCCCTC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 TTTAGTGAAGAACTCTTGATGAAGATCCAACTATCAACTTAAAGACGCTGATATTGTC 357
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrValLys 120
Db 358 AATGAAGTCAAGGTGTTTATATCATCATGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGT 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGACGCTCATGCTGATATGTTTCAAACTAAAGATCAAAATCAATCGTCAAAAAACAAG 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAArgAlaGlnGly 160
Db 478 CATGTCAAAGAT---AATGAGAAGTTAACTTAATGTTCTGTAGCAAGGCTCTCAGGGA 534
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACGACAAATGATGCTTATGCTTTAATCCACCTGATATTATCGAAGATACGGGT 594
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTTCCTCATGAGGTCACTATCACTTACATTCCTCAAAAGCGATTATCT 654
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg 218
Db 655 GCTAGTGAATTAGCAGCAGCTAGTGAACATATCGTGGTGAAGAAATAATGCAACCGAGTCAG 714
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGTATTCTTCAACAGCTAGTGAACATATAACACGCAA-----TCTGTAGCAAAA 765
QY 239 AsnLeuThrValThrProThrTyrHisGlnAspGlnGlyGluAsnIleSerSerLeuLeu 258
Db 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG 816
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 817 AAGGAACCTCTATGATTCACCTAGCGCCCAACGTTACAGTGAATCAGATGCGCTGCTTT 876
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCTGCTAAGATTATCATGCTGACCAAAATGGAGTTGCGATTCCGCATGGCGCAT 936
QY 299 TyrHisPheIleProTyrGlnMetSerGluLeuGluLysArgIleAlaArgIleIle 318
Db 936
```

```

937 TACCACTTTATTCCTTACAGCAAGCTTTCTGCTTAGAAGAAAGATTGCCAATGGTG 996
938 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338
939 CCTATCAGTGAAGTGGTCTCAGTGTCTTACAAATGCAAAACCTAAT----- 1044
939 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
1045 -----GAAGTAGTGTCTAGTCTAGGCACTCTTTCA 1074
359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
1075 AGCAATCCTCTCTTTACAGCAAGTAGGAG---CTCTCTCAGCATCTGATGTTAT 1131
379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
1132 ATTTTAAAT-----CCAAAGATATCGTTGAAGAACG 1164
399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysLeuSerHisLysLeuGlyAla 418
1165 GCTACA-----GCTTATATTGTAAGACATGGTGCATTCATTCCATTAC----- 1206
419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
1207 -----ATTCCAAATCAAAAT----- 1221
439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
1222 ---CAAATGGCAACCGACTCTTCCAAACATAGTCTAGCAACACTTCTCCATCTCT 1278
459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
1278 ----- 1278
479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
1279 -----CCAAATC-----AATCCAGGA 1293
499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGlu 518
1294 ACTTCACATGAGAAACATGAA-----GAA 1317
519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
1318 GATGATACGGATTTGATGCTAATCGTATTATTCGCTGAAGATGAATCAGGTTTGTGATG 1377
539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLysLeuGluAlaGlu 556
1378 AGTCACGGAGACCAATCATTTCTTCAAGAGGACTTGACAGAGAGCAA 1431

```

RESULT 28

```

US-09-765-272-181
: Sequence 181, Application US/09765272
: Patent No. US20020061545A1
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/765,272
: FILING DATE: 22-Jan-2001

```

```

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 181:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 181:
US-09-765-272-181

Alignment Scores:
Pred. No.: 9,36e-94 Length: 1342
Score: 1141.50 Matches: 250
Percent Similarity: 59.41% Conservative: 75
Best Local Similarity: 45.70% Mismatches: 119
Query Match: 28.40% Indels: 103
DB: 9 Gaps: 14

US-09-765-272A-66 (1-763) x US-09-765-272-181 (1-1342)
QY 5 LeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSerTyrIleAsp 24
Db 2 CTAACCCAGCATCGTTCG---CAGGAAATAAGGACAATAATCGTCTCTTATGTGGAT 58
QY 25 GlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSerLysArgGlu 44
Db 59 GGCAGCCAGTCAAGTCAAGAAAGTGAACCACTTGACACAGACCCAGGTTAGCCAGAAAGAA 118
QY 45 GlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyrValThrSerHis 64
Db 119 GGAATTCAGGCTGAGCAAAATTTGTAATCAAAATTCAGATCAGGGCTATGTACGTCACAC 178
QY 65 GlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIleIleSerGluGlu 84
Db 179 GGTGACCATCATCTACTATAATGGGAAAGTCTCTTATGATGCCCTCTTATGTGAAGAA 238
QY 85 LeuLeuMetLysAspProAsnTyrGlnLysLysAspSerAspIleValAsnGluIleLys 104
Db 239 CTCTTGATGAAGGATCCAAACTATCACTTAAGACGCTGATATTGTCAATGAAGTCAAG 298
QY 105 GlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHis 124
Db 299 GGTGGTTATATCATCAAGTGCATGGGAAATATTATGTTCTACCTGAAAGATCGACTCAT 358
QY 125 AlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnLysGlnLysArgSerHisAsn 144
Db 359 GCTCATATGTTCGAATCAAGATGAATCAATCGTCAAAACACAGACATGCTCAAGAT 418
QY 145 HisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrThr 164
Db 419 ---AATGAGAAGGTTAACTCTAATGTTGCTGTAGCAAGGTCTCAGGACGATATACGACA 475
QY 165 AspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIle 184
Db 476 AATGATGGTTATGTCTTTAATCCAGCTGATATTATCGAAGATACGGGTAATGCTTATATC 535
QY 185 ValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeu 204
Db 536 GTTCTCATGGAGGTCACTATCATCTACATTCCTCAAAAGGATTTATCTGCTAGTAGTAATTA 595
QY 205 AlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArgProSerSer 222
Db 596 GCAGCAGCTAAAGCACATCTCGCTGGGAAAAAATAATGCAACCGAGTCAGTTAAGCTATTCT 655

```


346 ATTATTAGTGAAGAGTTGTTGATGACGGATCCTAATTACCGTTTAAACAATCAGACGTT 405
 100 ValAsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrValTyrLeu 119
 406 ATCAATGAATCTTAGACGGTTACGTTTAAAGTCAATGGCACTATTATGTTACCTC 465
 120 LysAspAlaIleHisAlaAspAsnIleArgThrLysGluIleLysArgGln 137
 466 AAGCCAGTAGTAAGCGCAAAACATTCGAACCAACAACAATTCGTCAGCAAGTAGCC 525
 138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149
 526 AAAGGACTAAAGCAAGCTAAAGAAAAGTTAGCTCAAGTGGCCCATCTCAGTAAAGAA 585
 150 Asp---AsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyr 168
 586 GAAGTTTCGGCAGTCAATGAAGCAAAAAGACAAGACGCTATACACAGCATGGCTAT 645
 169 IlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValProHisGly 188
 646 ATTITTAGTCCGACAGATATCATTTGATGATTTAGGAGATGCTTATTAGTACCTCATG 705
 189 AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGlu 208
 706 AATCATTATCATTTATCTTAAAGAAATTTGCTCCAGTGAGCTAGCTGCTGCACAA 765
 209 AlaTyrTrpAsnGlyLysGln-----GlySerArgProSerSerSerSerTyrAsn 226
 766 GCCTACTGGAGTCAAAAACAAGTTCGAGGTCTAGACGCTGATTACCGCCGACCA 825
 227 Ala-----AsnProAlaGlnProArg 233
 826 GCCCCAGGTGCTAGAAAGCCCAATTCCTGATGTGACGCCCTAACCTGGACAAGTTCAT 885
 234 LeuSerGluAsnHisAsnLeuThrValThrPro-----Thr 245
 886 CAGCCAGATACGGTGGCTATCATCCAGCGCTCTAGCGCAAAATGATCGCTCAAAAAC 945
 246 TyrHisGlnAsn-----GlnGlyLysAsnIleSerSerLeuLeuArgGluLeuTyr 262
 946 AAACACCAAGACATGAGTTTAAAGGAAAACCTTTAAGAACCTTTTAGATCAACTACAC 1005
 263 AlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282
 1006 CGCTTTGATTGAATATACCTCATGTGGAGAAGATGGTTGATTTTGAACCACTACAC 1065
 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302
 1066 GTGATCAAAATCAACCGCTTTTGGGTATGTGGTGCCTCATGGAGATCATTTATCATTTATC 1125
 303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAla---ArgIleIleProLeuArg 321
 1126 CCAAGAGTCAGTTATCACTCTTGAATGGAAATTAGCATCGATCATCTTAGCTGGCCAA 1185
 322 TyrArgSerAsnHisTrpValProAspSerArgProGlu-----GlnProSerProGln 339
 1186 ACTGAGGACAAT-----GACTCAGGTTCAAGACACTCAAAACCATCATGATAA 1233
 340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359
 1234 GAAGTGACATACCTTTCTTGGTCTATCGCATCAAAAGCTTACGGAAAAGGCTTAGATGGT 1293
 360 AsnProIleAspGluLys-----LeuValLysGluAlaValArgLysVal--- 374
 1294 AAACCATATGATACGAGTATGCTTATGTTTATGTTTAAAGAAATCCATTCATTACGTGGAT 1353
 375 -----GlyAspGlyTyr----- 378
 1354 AAATCAGGAGTTACAGCTAAACACGAGATCATTTCCACTATATAGATTTGGAGAACTT 1413
 379 ---ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGlu 397
 1414 GAACAAATATGATGAGTTCGAGTTCGTAACCTGGTGAAGCAAAAGGTCAAGCTGATGAG 1473

QY 398 ThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeu-----SerHisLys 415
 Db 1474 CTTGCTGCTGCTTTGGATCAGGAACAGGCAAGAAAACACACTCTTTTGACATAAAAAA 1533
 QY 416 LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArg 428
 Db 1534 GTGAGTCGCAAAAGTAACAAAAGATGTAAGTGGCTATATGATGCCAAAAGATGGTAAG 1593
 QY 429 GluPheTyr-----AsnLysAlaTyrAspLeu-----LeuAlaArgIleHisGlnAsp 444
 Db 1594 GACTATTTCTATGCTGCTGATCACTCACTTGAATTCAGTTCAGATTGCTTTGCCGCAACAAGAA 1653
 QY 445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGlu--- 463
 Db 1654 CTAATGCTTAAAGATAAAGACGATTACCGTTATGACATTGTTGACACACAGGTATTGAGCCA 1713
 QY 464 ArgLeuLys---AspVal**Ser----- 470
 Db 1714 CGACTTGTGTAGATGTGTCAAGTCTGCCGATCGATGCTGTGTAATGCTACTTACGATACT 1773
 QY 471 -----AspLysValLysLeuVal----- 476
 Db 1774 GGAAGTTCTGTTTATCCACATATTTGATCATATCCATCTGCTTCGTTTCGTTATCATGTTG 1833
 QY 477 -----***AspIleLeuAlaPheLeuAlaProIleArgHisProGlu----- 490
 Db 1834 ACGCGGATCAGATTGCAACAGTCAAGTATGTCATGATGCAACACCCCGAAGTTTCGTCGGAT 1893
 QY 490 ----- 490
 Db 1894 GTATGCTTAAGCCAGGCGCATGAAGAGTCAGTTTCGTCGTCATTCCTCAATGTTACGCTCTT 1953
 QY 491 -----ArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVal 508
 Db 1954 GATAAACGTCGTGTTATGCAAACTGCAAAATTTATCCATTCCTCTGAAGAAGTTCAAAA 2013
 QY 509 AlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArgAspIle 528
 Db 2014 GCCCTAGCAGAGGTCGTTTTCACACACGACGCGTATATTTTCGATCCACGAGATGTT 2073
 QY 529 ThrSerAspGlu-----GlyAspAlaTyrValThrProHisMetThrHis 543
 Db 2074 TTGCCAAAGAAACTTTTGTATGGAAGATGCTCTTTAGCATCCCAAGACGATGGC 2133
 QY 544 Ser-----HisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln 561
 Db 2134 AGTTCATTGAGAACCATTAATAAATCTGATCTATCCCAAGCTGAGTGGCAACAGTCAA 2193
 QY 562 AlaTyrAlaLysGluLysGlyLeu-----ThrProProSerThr 574
 Db 2194 GAGTTATTGGCAAGAAAATACTGGTATGCTACTGATACGGATAAACCCCAAGAAAG 2253
 QY 575 AspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVal 594
 Db 2254 CAACAGCAGATAGACGATGAACCAACCAAGCCAGTGAAGCC-----AGTAAAGAA 2307
 QY 595 LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614
 Db 2308 GAAAAAGAAATCAGATGACTTTATAGACAGTTTACCA----- 2343
 QY 615 ValLysAsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGlu 634
 Db 2343 ----- 2343
 QY 635 TrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAla 654
 Db 2344 -----CACTATGCTCTAGATAGACCA-----ACCCTAGAAGATCATATCAAT 2385
 QY 655 ThrVal-----LysTyrTyrValGluHisProAsnGluArg 666
 Db 2386 CAATTAGCAAAAAGCTAATATATCGATCCTAAGTATCTCTATTTCCTCAACCA----- 2436

QY 667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGln---ArgAsnLysAsn 695
 Db 2437 -----GAAGGTGTCCTCAATTTTATATAATAAATAAT 2463

QY 686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
 Db 2464 GGTGAATTGGTAACCTATGATAT-CAAGACACTTCACAAATAAACCCCTAACCAAAAGA 2522

RESULT 30

US-10-340-792-13
 ; Sequence 13, Application US/10340792
 ; Publication No. US2003022832A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BRODEUR, Bernard R.
 ; APPLICANT: RIOUX, Clement
 ; APPLICANT: BOVER, Martine
 ; APPLICANT: CHARLEBOIS, Isabelle
 ; APPLICANT: HAMEL, Josee
 ; APPLICANT: MARTIN, Denis
 ; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
 ; FILE REFERENCE: 8331-9002
 ; CURRENT APPLICATION NUMBER: US/10/340,792
 ; PRIOR FILING DATE: 2003-01-13
 ; PRIOR APPLICATION NUMBER: US/09/252,088B
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US/60/075,425
 ; PRIOR FILING DATE: 1998-02-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 5215
 ; TYPE: DNA

ORGANISM: group B streptococcus

FEATURE:

NAME/KEY: CDS

LOCATION: (3) .. (122)

FEATURE:

NAME/KEY: CDS

LOCATION: (133) .. (2511)

FEATURE:

NAME/KEY: CDS

LOCATION: (367) .. (2511)

FEATURE:

NAME/KEY: CDS

LOCATION: Complement ((2716) .. (2946))

FEATURE:

NAME/KEY: CDS

LOCATION: Complement ((2295) .. (3252))

FEATURE:

NAME/KEY: CDS

LOCATION: Complement ((3299) .. (3676))

FEATURE:

NAME/KEY: CDS

LOCATION: Complement ((3837) .. (4124))

FEATURE:

NAME/KEY: CDS

LOCATION: Complement ((4351) .. (5214))

FEATURE:

US-10-340-792-13

Alignment Scores:

Pred. No.: 2,96e-82 Length: 5215

Score: 1023.50 Matches: 289

Percent Similarity: 46.05% Conservative: 107

Best Local Similarity: 33.60% Mismatches: 253

Query Match: 25.47% Indels: 212

DB: 16 Gaps: 33

US-09-765-272a-66 (1-763) x US-10-340-792-13 (1-5215)

QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21

Db 109 AGTTACCACTTGTGAAGCATCATATGGGTGTAGCAAAAGGAC---AATCAGATTGCC 165

QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGlu 39
 Db 166 TATATTGATGACGCAAAAGGTAAAGCAAAAGCCCTTAAACAAAACAAACGATGATCAA 225
 QY 40 ValSerIysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGly 59
 Db 226 ATCAGTGTCTGAAGAAGCATCTCTGCTGAACAGATCGTAGTCAAAATTTACTGACCAAGGC 285
 QY 60 TyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProThrAspAla 79
 Db 286 TATGTGACCTCACAGGTGACCATATCATTTTACATGGGAAAGTTCTTATGATGCG 345
 QY 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysLysAspSerAspIle 99
 Db 346 ATTATTAGTCAGAGTGTGTTGATGACGGATCTTAATTTACCGTTTAAACAATCAGACGT 405
 QY 100 ValAsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrValTyrIleu 119
 Db 406 ATCAATGAATCTTAGACGGTTAGCTTATTAAAGTCAATGGCAACTATTATGTTTACCTC 465
 QY 120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGln----- 137
 Db 466 AAGCCAGGTAGTAAGCCAAACATTCGAACCAACAAATTTGCTGACCAAGTAGCC 525
 QY 138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149
 Db 526 AAAGGAACCTAAAGAAAGCTAAAGAAAGGTTTAGCTCAAGTGGCCCATCTCAGTAAAGAA 585
 QY 150 Asp---AsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyr 168
 Db 586 GAAGTTGCGCGCAGTCAATGAAGCAAAAGCAAGGACGCTATCTACAGACGATGGCTAT 645
 QY 169 IlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValProHisGly 188
 Db 646 ATTTTGTAGTCCGACATATCATGATGATGATGATGATGATGATGATGATGATGATGATG 705
 QY 189 AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGlu 208
 Db 706 AATCACTATCATATATATCTCTAAAGAGGATTTGCTCCAAGTGAAGTGAAGTGAAGTGA 765
 QY 209 AlaTyrTyrAsnGlyLysGln-----GlySerArgProSerSerSerSerSerTyrAsn 226
 Db 766 GCCTACTGGAGTCAAAACAAAGGTCGAGGTGCTAGACCGTCTGATTACCGCCGACACCA 825
 QY 227 Ala-----AsnProAlaGlnProArg 233
 Db 826 GCCCCAGGTGCTGAGAAAGCCCAATTCCTGATGTGAGCGCTAACCCCTGGACACGTCAT 885
 QY 234 LeuSerGluAsnHisAsnLeuThrValThrPro-----Thr 245
 Db 886 CAGCCAGATAACGGTGGCTATCATCCAGCGCTCTTAGGCCAAATGATGGTGCACAAAC 945
 QY 246 TyrHisGlnAsn-----GlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyr 262
 Db 946 AAACACCAAGAGATGAGTTTAAAGGAAAAACCTTTAAGGAACCTTTTAGACTCAACTAC 1005
 QY 263 AlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282
 Db 1006 CGTCTTGATTGAATATACCGTATGGAAGAGATGGGTGTTGATTTTGAACCGACTCAA 1065
 QY 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302
 Db 1066 GTGATCAAAATCAACGCTTTTGGGTATGTGGTGCCTCATGGAGATCATATATCATATATC 1125
 QY 303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAla---ArgIleIleProLeuArg 321
 Db 1126 CCAAGAAGTCAGTTATCATCTCTGAAATGGAATAGCAGATCGATCTAGCTGGCCAA 1185
 QY 322 TyrArgSerAsnHisTyrValProAspSerArgProGlu-----GlnProSerProGln 339
 Db 1186 ACTGAGGCAAT-----GACTCAGGTTTCAGACACTCAAAACCATCAGATAA 1233
 QY 340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359

```
1234 GAAGTACACATACCTTTCTTGGTCTACGATCAAGCTTAAAGAGGCTTAGATGGT 1293
      ::::
360 AspProLeuAspGluLys-----LeuValLysGluAlaValArgLysVal--- 374
      |||
1294 AAACCATATGATACGAGTGATGCTTATGTTTATGTAAGATCAATTCATTCAGTGGAT 1353
      |||
375 -----GlyAspGlyTyr----- 378
      ::
1354 AAATCAGGAGTTACAGCTAAACACGAGATCAATTCACCTATATAGGATTTGGAGAACTT 1413
      ::
379 ---ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGlu 397
      ::::
1414 GAACATATGAGTTGATGAGTGGTCTTAAGTCTGAGTGAACAAAGCTCAAGCTGATGAG 1473
      ::
398 ThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeu-----SerHisLys 415
      |||
1474 CTTGCTGCTCTTGGATCAGGACAGGACAGGCAAGAAACACCTCTTTGACACTAATAAAA 1533
      ::
416 LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArg 428
      ::
1534 GTGAGTCGCAAGATTAACAAAGATGCTAAAGTGGCTATATGATGCCAAAGATGGTAAG 1593
      ::
429 GluPheTyr-----AsnLysAlaTyrAspLeu-----LeuAlaArgIleHisGlnAsp 444
      ::
1594 GACTATTTCTATGCTCGTGCATCAACTTGATTTGACTCAGATTCCTTTGCGCAACAGAA 1653
      ::
445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGlu--- 463
      |||
1654 CTAATGCTTAAGATGATGAGCAATTAACGCTTATGACATGTTGACACAGGATTTAGCCCA 1713
      ::
464 ArgLeuLys---AspVal***Ser-----AspLysValLysLeuVal----- 470
      |||
1714 CGACTTGTGTAGATGTGTCAAGTCTGCCATGCTGCTGTAATGCTACTTACGATACT 1773
      ::
471 -----AspLysValLysLeuVal----- 476
      |||
1774 GGAAGTTCGTTTGTATCCCATATTTGATCATCATCATGCTGCTCGTATTTCATGCTTG 1833
      ::
477 -----AspIleLeuAlaPheLeuAlaProIleArgHisProGlu----- 490
      |||
1834 AGCGCGCATCAGATTGCAACAGTCAAGTATGTGATGCAACACCCCGAAGTTCGTCGGAT 1893
      ::
490 ----- 490
      ::
1894 GTATGCTTAAGCCAGGCGATGAAGAGTTCAGTTCGTTCCATTCCTCAATGTTACGCTCTT 1953
      ::
491 -----ArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVal 508
      |||
1954 GATAAACGTGCTGGTATGCCAAACTGGCAAAATATCAATTCCTGCTGAAGAAGTTCAAAA 2013
      ::
509 AlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProArgAspIle 528
      |||
2014 GCCCTAGCAGAAAGGTCGTTTTCGAACACACAGCGGTATATTTTCGATCCACGAGATGT 2073
      ::
529 ThrSerAspGlu-----GlyAspAlaTyrValThrProHisMetThrHis 543
      ::
2074 TTGGCAAGAAACTTTTGTATGGAAGATGCTCTTTAGCATCCCAAGACAGATGGC 2133
      ::
544 Ser-----HisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln 561
      |||
2134 AGTTCATTGAGAACCATTAATAATCTGATCTATCCAGCTGAGTGGCAACAGCTCAA 2193
      ::
562 AlaTyrAlaLysGluLysGlyLeu-----ThrProProSerThr 574
      ::
2194 GAGTTATGGCAAGAAATAATACTGGTGTGCTACTGATPACGATATAACCCCAAGAAAG 2253
      ::
575 AspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVal 594
      |||
2254 CAACAGGAGATAGACGATGAAACCAACAGCCAGTGAAGGCC-----AGTAAAGAA 2307
      ::
595 LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614
      ::
```

```
Db 2308 GAAAAGAATACAGATGACTTTTATAGACAGCTTTACCA----- 2343
QY 615 ValLysAsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGlu 634
Db 2343 ----- 2343
QY 635 TrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAla 654
      |||
Db 2344 -----GACTATGCTGTAGATAGCA-----ACCTAGAGATCATATCAAT 2385
      |||
QY 655 ThrVal-----LysTyrTyrValGluHisProAsnGluArg 666
      ::
Db 2386 CAATTAGCACAAAAAGCTAATATCGATCTCAAGTATCTCTATTTTCCAACCA----- 2436
      |||
QY 667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGln---ArgAsnLysAsn 685
      |||
Db 2437 -----GAAGGTGTCCAAATTTTATATAATAAAT 2463
      |||
QY 686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
      |||
Db 2464 GGTGAATTGGTAACTTATGATAT-CAAGACACTTCAACAATAAACCTTACCAAAAGA 2522
      |||

RESULT 31
US-09-769-736-17
; Sequence 17, Application US/09769736
; Publication No. US20030138775A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089wo
; CURRENT APPLICATION NUMBER: US/09/769,736
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-09-769-736-17

Alignment Scores:
Pred. No.: 2,46e-82 Length: 2469
Score: 1019.50 Matches: 282
Percent Similarity: 44.52% Conservative: 100
Best Local Similarity: 32.87% Mismatches: 239
Query Match: 25.37% Indels: 237
DB: 10 Gaps: 32

US-09-765-272a-66 (1-763) x US-09-769-736-17 (1-2469)
QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21
      |||
Db 64 AGTTACAGCTTGGTAAAGCATCATATGGTCTAGCAACAAGGAC---AATCAGATTGCC 120
      |||
QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGlu 39
      |||
Db 121 TATATTGATGATAGCAAGAGTAAAGGTAAAGCCCTAAACCAACAAACACGATGGATCAA 180
      |||
QY 40 ValSerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGly 59
      |||
Db 181 ATCAGTCTGAAGAGGACATCTCTGCTGAACAGATCTGCTGAAATTTACTGACCAAGGT 240
      |||
QY 60 TyrValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAla 79
      |||
Db 241 TATGTTACCTCACACGGTGAACCATATATTTTACATGCGGAAGTTCCTTATGATGCG 300
      |||
```

QY 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIle 99
Db 301 ATTATTAGTGAAGAGTCTGTTGATGACGGATCTTAATTACCAATTTTAAAAACAATCAACAGCTT 360
QY 100 ValAsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeu 119
Db 361 ATCAATGAACTTAGACGGTTACGTTATTAAAGTCATGGAACATTAATGTTTACCTC 420
QY 120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGln 137
Db 421 AAGCCAGGTAGTAAGCGCAAAACATTCGAACCAACAAACAAATGCTGAGCAAGTAGCC 480
QY 138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149
Db 481 AAGGAACCTAAAGAGTAAAGAAAAAGGTTTAGCTCAAGTGGCCCACTCTCAGTAAAGAA 540
QY 150 Asp---AsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyr 168
Db 541 GAAGTTGGCGGAGTCATGAAGCAAAAGACAAAGACGCTATATACAGCATGGCTAT 600
QY 169 IlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValProHisGly 188
Db 601 ATTATTTAGTCGACAGATATCATTTGATGATTTAGGAGATGCTTATTAGTACCTCATGTT 660
QY 189 AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGlu 208
Db 661 AATCACTATCATTTATTTCTTAAAAAAGATTGTTCTCAAGTAGAGTAGCTGCTGCACAA 720
QY 209 AlaTyrTrpAsnGlyLysGln-----GlySerArgProSerSerSerSerTyrAsn 226
Db 721 GCCTACTGGAGTCAAAACAAAGTCGAGGTGCTAGACCGTCTGATTACCGCCGACACCA 780
QY 227 Ala-----AsnProAlaGlnProArg 233
Db 781 GCCCCAGGTCGTAGAAAAGCCCCAATTCCTGATGTGACGCTTAACCTTGGCAAGGTAT 840
QY 234 LeuSerGluAsnHisAsnLeuThrValThrPro-----Thr 245
Db 841 CAGCCAGATACGGTGGTTATCATCCAGCGCTCTAGGCCAAATGATCGCTCACAAAAC 900
QY 246 TyrHisGlnAsn-----GlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyr 262
Db 901 AAACACAAAGAGATGATTTAAAGGAAAAACCTTTAAGGAACCTTTAGATCATCTACAC 960
QY 263 AlalysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282
Db 961 CGCTCTGATTGAATACCGTCATGTGGAAGAAGATGGTTGATTTTGAACCGACTCAA 1020
QY 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302
Db 1021 GTGATCAAAATCAAAACGGCTTTTGGGTATGTGTGCTCATGGAGATCATATCATATTATC 1080
QY 303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyr 322
Db 1081 CCAAGAGTCAGTTATCACTCTTGAATGAATAGAGAT-----CGATAC 1128
QY 323 ArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSerThrPro 342
Db 1129 TTAGCCGCGC-----CAAAGTATGATC 1149
QY 343 GluProSerProSerProGlnProAlaProAsnProGlnProAla----- 357
Db 1150 AAGGACTCAGGTTTCAGATCACTCAAAACCAATCAGATAAAGAAAGTGACACATACCTTTCTT 1209
QY 358 -----ProSerAsnProIleAspGluLys--- 365
Db 1210 GGTATCGCATCAAAAGCTTACGAAAAAGGCTTAGATGGTAAACCATATGATACAGAGTAT 1269
QY 366 -----LeuValLysGluAlaValArgLysVal----- 374
Db 1270 GCTATGTTTTTAGTAAGAATCCATTCATTCAGTGGATAAATCAGGAGTTTACAGCTAAA 1329
QY 375 ---GlyAspGlyTyr-----ValPheGluGluAsnGly 384

Db 1330 CACGGAGATCATTTCCACTATATAGGATTTGGAGAACTTGAACAATATGATGCTGATGAG 1389
QY 385 ValSerArgTyrIleProAlaLysAspLeuSerAlaGlnThrAlaAlaGlyIleAspSer 404
Db 1390 GTCCGTAACCTGGGTGAAGCAAAAGGTCAAGCTGATGAGCTTGTGCTTGGATCAG 1449
QY 405 LysLeuAlaLysGlnGluSerLeu-----SerHisLysLeuGlyAlaLysLysThrAsp 422
Db 1450 GAACAAGGCAAGAAAAACCACTCTTTGACACTAAAAAAGTGAAGTCCCAAGTAAACAAA 1509
QY 423 -----LeuProSerSerAspArgGluPheTyrAsnLysAlaTyr 435
Db 1510 GATGTTAAAGTGGGCTATATATATGCAAAAGATGGCAAGGACTATTCTATGCTCGTTAT 1569
QY 436 AspLeu-----LeuAlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArg 451
Db 1570 CAACTTGATTTGACTCAGATTGCTTTGGCAACAAAGAACTAATGCTTAAAGATAAGAA 1629
QY 452 GlnValAspPheGluAlaLeuAspAsnLeuGlu-----ArgLeuLys---AspVal*** 469
Db 1630 CATTACCGTTATGACATCTGTTGATACAGGCATTGAGCCAGACTTGTCTGTAGATGTCA 1689
QY 470 Ser----- 470
Db 1690 AGTCTGCCGATGCTGCTGTAATGCTTACTGATGCTGGAAGTTCGTTTGTGTTATCCCA 1749
QY 471 -----AspLysValLysLeuVal-----***AspIleLeuAla 481
Db 1750 CATATTGATCATATCCATGCTGCTTCGTTATTCATGTTGACGCAATCAGATTGCAACA 1809
QY 482 PheLeuAlaProIleArgHisProGlu----- 490
Db 1810 ATCAAGTATGTGATGCAACACCCCGAAGTTCGTCGCGATGTATGTTTAAGCCAGGCGAT 1869
QY 491 -----ArgLeuGlyLysPro 495
Db 1870 GAAGAGTCAGGTTCCGTCATTCATAATGTTAGCCCTCTTGATAAAGCTGCTGGTATGCCA 1929
QY 496 AsnAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyr 515
Db 1930 AACTGGCAATTTATCCATTTCTGCTGAAGAAGTTCAAAAAGCCCTAGCAGAGGTCGTTTT 1989
QY 516 ThrThrGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGlu----- 532
Db 1990 GCAGCACAGACGGCTATATTTTGCATCCAGAGATGTTTGGCAAAAGAAACTTTTGTGA 2049
QY 533 -----GlyAspAlaTyrValThrProHisMetThrHisSer-----HisTrpIleLys 548
Db 2050 TGCAAAAGATGGCTCCTTTAGCATCCCAAGACAGATGGCAGTTTCATTGAGAACCATTAA 2109
QY 549 LysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGly 568
Db 2110 AAATCCGATCTATCCCAAGCTGAGTGGCAACAAAGCTCAAGATTATGGCAAGAAATAAT 2169
QY 569 Leu-----ThrProProSerThrAspHisMetThrHisSer-----HisTrpIleLys 581
Db 2170 GCTGTGTGCTGCTGATACGGATAAATCTGAAGAAAGCAACAGCAGCATAGAGCAAT 2229
QY 582 ThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaLysLysValPro 601
Db 2230 GAAAAACCAACAGCCCAAGTGAAGCC-----AGTAAAGAAAGAAAGAAATCAGATGACTTT 2283
QY 602 LeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIle 621
Db 2284 ATAGACAGTTTACCA----- 2298
QY 622 IleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyr 641
Db 2299 -----GACTATGCTGTAGAT 2313
QY 642 GluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrVal----- 656

Db 2314 AGAGCA-----ACCTAGAGATCATATCAATCAATAGCACAAAAAGCTAAT 2361
 QY 657 -----LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPhe 673
 Db 2362 ATCGATCCTAAGTATCTCAATTTCCAAACCA-----2391
 QY 674 GlyAsnAlaSerAspHisValGln---ArgAsnLysAsnGlyGlnAlaAspThr 690
 Db 2392 -----GAAGGTGTCCCAATTTTATAATAAAAAATGGTGAATTGGTAAC 2433

RESULT 32

US-09-884-465A-9
 ; Sequence 9, Application US/09884465A
 ; Publication No. US20030077293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shire Biochem, Inc.
 ; APPLICANT: Hamel, Josee
 ; APPLICANT: Brodeur, Bernard
 ; APPLICANT: Martin, Denis
 ; APPLICANT: Charland, Nathalie
 ; APPLICANT: Ouellet, Catherine
 ; TITLE OF INVENTION: Streptococcus Antigens
 ; FILE REFERENCE: 055190-0044
 ; CURRENT APPLICATION NUMBER: US/09/884,465A
 ; CURRENT FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: 60/212,683
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 384
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 2528
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-09-884-465A-9

Alignment Scores:

Aligned. No.: 9,64e-79 Length: 2528
 Score: 980.00 Matches: 282
 Percent Similarity: 44.50% Conservative: 127
 Best Local Similarity: 30.69% Mismatches: 228
 Query Match: 24.38% Indels: 282
 DB: 10 Gaps: 36

US-09-765-272A-66 (1-763) x US-09-884-465A-9 (1-2528)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
 Db 1 TTGGCTATGCCTAACCAACGATCGTTCG---CAGGAAATAGGACAAATATCGGTGC 57
 QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
 Db 58 TCTTATGTGATGGCAGCCAGTCAAGTCAGAAAGTGAANAATTTGACACAGACAGGTT 117
 QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
 Db 118 AGCCGAAAGAGGAATTCAGGCTGAGCAATTTCTAATCAAAATTAACATCAGGCTAT 177
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
 Db 178 GTAACGTACACGGTGATCACTATCAATTAATATGGAAGTTCCTTATGATGCCCTC 237
 QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysLysAspSerAspIleVal 100
 Db 238 TTTAGTGAGAACTCTTGATGAGAGATCCAACTATCACTTAAGACGCTGATTTGTC 297
 QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
 Db 298 AATGAAGTCAAGGGTGTATATCATCAAGTCAGTGAAGTGAAGTATATGCTACCTGAAA 357
 QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
 Db 358 GATCGAGCTCATGCTGATAATGTTCCAACTAAAGATCAATCAATCGTCAAAAAAGAA 417

QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
 Db 418 CATGTCAAAAGT---AATGAGAAGGTAACTCTAATGTTGCTAGCAAGGTCTCAGGGA 474
 QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
 Db 475 CGATATACGACAAATGATGGTTATGTCCTTAATCCAGCTGATATATCGAAGATACCGGT 534
 QY 181 AspAlaIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
 Db 535 AATGCTTATATCGTCTCATGAGGTCACTACTACATCCCAAAAGCGATTATCT 594
 QY 201 AlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220
 Db 595 GCTAGTGAATTAGCAGCAGCTAAGCACATCTCGCTGGAATA---AATATGCAACCG 648
 QY 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240
 Db 649 AGTCAGTTAAGCTATTCTTCAACACCTCTCCA---TCTCTT 687
 QY 241 ThrValThrPro---ThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259
 Db 688 CCAATCAATCCAGGAACCTTCACAT-----711
 QY 260 GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279
 Db 712 -----GAGAAACAT---GAAGAAGATGGATACCGATTGAT 744
 QY 280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299
 Db 745 GCTAATCGTATTATCGCTGAAGATGAATCAGGTTTGTGTCATGATCAGTCACGAGACCAAT 804
 QY 300 HisPhe-----IleProTyrGluGlnMetSerGluLeuGluLysArgIle 314
 Db 805 CATTAATCTTCAAGAAGGACTTGCAGAGAAGCAAAATTAAGCTCGCGAAAAACATTTA 864
 QY 315 AlaArgIleIleProLeuArgTyrArgSerAsnHisTyrValProAsp-----SerArg 332
 Db 865 GAGGAAGTT-----AAACTAGTCATAATGGATTAGATTCTTTGTCTATCT 909
 QY 333 ProGluGlnProSerProGlnSerThrProGluProSerProSerProGlnProAla--- 351
 Db 910 CATGAACAGGATTAATCCAAAGTAATGCCAAAGAAATGAAGATTTAGATAAAAAATCGAA 969
 QY 351 -----351
 Db 970 GAAAAATGTGTCATTATGAAACAAATATGTTGTCGTAACAGTAAAGTATGTCGTAAT 1029
 QY 352 -----ProAsnProGlnProAlaProSerAsnProIleAsp 363
 Db 1030 AAGAAAAAATGCGATTATTATTCGCGATGGAGATCACCATCATGCGATCGATTGAT 1089
 QY 364 Glu-----364
 Db 1090 GAACATAAACCGTTGGAATTTGGTTCATTCTCACAGTAACATGTAAGTGTTTAAACCCGAA 1149
 QY 364 -----364
 Db 1150 GAAGAGTTGCTAAAAAAGAGGGAATAAGTTTATCTGGAGAGAAATTAACGAATGTT 1209
 QY 365 ---LysLeuValLys-----368
 Db 1210 GTTAATTTGTTAAAAAATAGTACGTTTAATTAATCAAACTTTACTTACCCAAATGTCAA 1269
 QY 369 -----GluAlaValArgLysValGly-----375
 Db 1270 AAACGCGTTCTTTTAGTTTCCGCTGAATTTGGAGAAAAAATTAGGTATCAATATGCTA 1329
 QY 376 -----AspGlyTyrValPheGluGlu-----382
 Db 1330 GTAAAAATTAATACACAGATGGAAGATTAATTTGGAGAAAGTATCTGGTAAAGTATTTGGA 1389
 QY 383 AsnGlyValSerArg-----TyrIleProAlaLysAsp 393

Db 2242 CCAGCAGATTCTTTTACCAGAGGCCACCAACGAAAAAACCCTGTAAAAACCGAGAAAACTCAACG 2303
 QY 720 -----LysProGluSerProLysProThrGluGluProGluGluSerPro 734
 Db 2302 GATAATGGAAATGTTGAATCCAGAAAGGAATGTGGGGAGTGACCCCTATGTTTAGATTCAAGCA 2361
 QY 735 GluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGlu 753
 Db 2362 TTAGAG--GAAGCTCCAGCAGTA-----GATCCCTGTACAGAAAAATTAGAAAAA 2409
 RESULT 33
 US-09-769-736-23
 ; Sequence 23, Application US/09769736
 ; Publication No. US20030138775A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Microbial Technics Limited
 ; APPLICANT: Le Page, Richard WF
 ; APPLICANT: Wells, Jeremy M
 ; APPLICANT: Hanniffy, Sean B
 ; TITLE OF INVENTION: Proteins
 ; FILE REFERENCE: PNC/P21089wo
 ; CURRENT APPLICATION NUMBER: US/09/769,736
 ; CURRENT FILING DATE: 2003-02-14
 ; PRIOR APPLICATION NUMBER: GB 9816335.5
 ; PRIOR FILING DATE: 1998-07-27
 ; PRIOR APPLICATION NUMBER: US 60/125163
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 23
 ; LENGTH: 1146
 ; TYPE: DNA
 ; ORGANISM: Streptococcus agalactiae
 US-09-769-736-23
 Alignment Scores:
 Pred. No.: 7.09e-70 Length: 1146
 Score: 877.00 Matches: 185
 Percent Similarity: 62.95% Conservative: 41
 Best Local Similarity: 51.53% Mismatches: 81
 Query Match: 21.82% Indels: 52
 DB: 10 Gaps: 11
 US-09-765-272A-66 (1-763) x US-09-769-736-23 (1-1146)
 QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21
 Db 64 AGTTACAGCTTGTTAAGCATCATATGGGTCTAGCAACAAGAGAC---AATCAGATTGCC 120
 QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGlu 39
 Db 121 TATATTGATGATAGCAAGTGAAGTAAAGCCCTTAAACCAACAAACGATGGATCAA 180
 QY 40 ValSerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGly 59
 Db 181 ATCAGTGTCTGAAGAGGCATCTCTGTGTAACAGATCGTAGTCAAAATTAAGTCAACCAAGT 240
 QY 60 TyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAla 79
 Db 241 TATGTTACTCTCACAGCGTGACCATATATCATTTTTTAAATGGGAAAGTCTCTTATGATGCG 300
 QY 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIle 99
 Db 301 ATTATTAGTGAAGAGTTGTGTATGACGATCTAATATTACATTTTAAACATCAGACGTT 360
 QY 100 ValAsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeu 119
 Db 361 ATCAATGAAATCTTAGACGGTTACGTTATTAAAGTCAATGGCAACTATTATGTTTACCTC 420
 QY 120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGln----- 137
 Db 421 AAGCCAGGTGTAAAGCGCGCAAAACATTCGAACCAACCAACAAATTCGTTCAGCAGTAGCC 480

138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149
Db AAAAGAACTAAAGAAAGAAAGAAAGTTAGCTCAAGTGGCCCACTCTCAGTAAAGAA 540
2Y 150 Asp---AsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyr 168
Db 541 GAAGTTCGGCAGTCAATGAAGCAAAAGAAAGCAAGAGCGCTATCTACAGAGATGGCTAT 600
169 IlePheAsnAlaSerAspIleGluAspThrGlyAspAlaTyrIleValProHisGly 188
Db 601 ATTTTGTAGTCGACAGATATCATGTAGTATTTAGAGATGCTTTATTTAGTACCTCATGCT 660
189 AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGlu 208
Db 661 AATCACTATCATTTATTTCTTAAAGAAAGATTGCTCAAGTGAAGTCTGCTGCACAA 720
209 AlaTyrTrpAsnGlyGlnGlySerArgProSerSerSerSerSerTyrAsnAlaSer 228
Db 721 GCCTACTGGAGTCAAAACCAAGGT---CGAGGTCTAGACCGTCTGATTACCGCCGACA 777
229 ProAlaGlnProArg-----LeuSerGluAsnHisAsnLeuThrValThrPro 244
Db 778 CCAAGCCCGAGTCTGTAGAAAGCCCACTCTCTGAT-----GTACGCGCT 822
245 Thr-----TyrHis----- 247
823 AACCTGGCAAGGTCTATCAGCCAGATAACGGTGGTTATCATCCAGCGCTCTTAGGCCA 882
248 -----GlnAsn-----GlnGlyGluAsnIleSerSer 256
Db 883 AATGATGCGTCAAAACCAAAACCAAGAGATGAGTTTAAAGGAAACCTTTAAGGAA 942
257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276
Db 943 CTTTATAGTCACTACCGCTCTTGATTTGAAATPACCGTCTATGCGGAGAGATGGTGG 1002
277 IlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGly 296
Db 1003 ATTTTGAACCGACTCAAGTATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1062
297 AsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAla 315
Db 1063 GATCATATCATATTTATCCCAAGAGTCACTTATCATCTCTTGAATGAATGAGTA 1119

RESULT 34
JS-09-884-465A-257
; Sequence 257, Application US/09884465A
; Publication No. US2003007293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josée
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIORITY FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIORITY FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
JS-09-884-465A-257

Alignment Scores: 1.37e-65 Length: 819
Pred. No.:

Score: 827.50 Matches: 169
Percent Similarity: 72.83% Conservative: 16
Best Local Similarity: 66.54% Mismatches: 28
Query Match: 20.59% Indels: 41
DB: 10 Gaps: 5
US-09-765-272A-66 (1-763) x US-09-884-465A-257 (1-819)
QY 498 GlnIleThrTyrAspAspGluIleGlnValAlaLysLeuAlaGlyTyrThrThr 517
Db 4 CAAATATACCTACATCATGATTCAGGTAGCCAAAGTTGGCAGCAAGTACACAA 63
QY 518 GluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVal 537
Db 64 GAAGACGGTTATCTTTGAT----- 84
QY 538 ThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArg 557
Db 85 -----ACTAGTTGGATTAAAGAGATAGTTTGTCTGAGCTGAGAGA 126
QY 558 AlaAlaAlaGlnAlaTyrAlaLysGlyLeuThrProProSerThrAspHisGln 577
Db 127 GCGGAGCCCGCTTATGCTTAAAGAGAAAGTTTACCCCTCTCTCGACAGACCAG 186
QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAla 597
Db 187 GATTTCAGGAATATCTAGGCAAGAGAGAGAGTATCTACCAACCGCTGAAAGCAGCT 246
QY 598 LysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsn 617
Db 247 AAGAGAGTGGCTATGATCGTATGCTTACATCTTCAGTATCTAGTACTAGTAAAGTCA 306
QY 618 GlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAsp 637
Db 307 GGTAGTTTAAATCATACCTCATTCATGACCATTAACATCAATCAATTTGAGTGTGTTGAC 366
QY 638 GluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLys 657
Db 367 GAAGCCCTTTATGAGGACCTTAAAGGGGTATAGTCTTGGAGATCTTTTGGCGACTGTCAAG 426
QY 658 TyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677
Db 427 TACTATGTGGA-----CGCGG-----AACGCTAGT 453
QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697
Db 454 GACCATGTTCTGTAATAATAAG-----GCAGACCAAGATAGTAAACCTGAT 498
QY 698 GluGlyLysProGlnThrGluLysProGluGluGluThrProArgGluGlyLysProGln 717
Db 499 GAAGATAGGAACATGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 558
QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731
Db 559 AATCAGCTGTTTAAATCTCTTCAGCAGATATCTTTTATAACCAAGCACTGATACGGA 618
QY 732 GluSerProGluGluSerGluGluProGlnValGluThrGlu 745
Db 619 GAGACAGAGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660

RESULT 35
US-10-324-143-10
; Sequence 10, Application US/10324143
; Publication No. US20030232976A1
; GENERAL INFORMATION:
; APPLICANT: HAMEL, JOSEE
; APPLICANT: CHARLAND, NATHALIE
; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: MARTIN, DENIS
; APPLICANT: BLAIS, NORMAND
; APPLICANT: OUELETTE, CATHERINE
; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 55190-56

; CURRENT APPLICATION NUMBER: US/10/324,143
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/341,252
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide sequence
US-10-324-143-10

Alignment Scores:
Pred. No.: 1,376-65 Length: 819
Score: 827.50 Matches: 169
Percent Similarity: 72.83% Conservative: 16
Best Local Similarity: 66.54% Mismatches: 28
Query Match: 20.59% Indels: 41
DB: 16 Gaps: 5

US-09-765-272A-66 (1-763) x US-10-324-143-10 (1-819)

QY 498 GlnIleThrTyThrAspGluLeuGlnValAlaLysLeuAlaGlyLysTyThrThr 517
Db 4 CAAATTTACCTACCTGATGAGATTCAGGTAGCCAGTTGGCAGGCAAGTACACAA 63

QY 518 GluAspGlyTyThrPheAspProArgAspIleThrSerAspGluGlyAspAlaTyVal 537
Db 64 GAAGACGGTTATATCTTTGAT 84

QY 538 ThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArg 557
Db 85 -----ACTAGTTGGATTAAAGAGATAGTTTGTCTGAAGCTGAGAGA 126

QY 558 AlaAlaAlaGlnAlaTyThrAlaLysGluLysGlyLeuThrProSerThrAspHisGln 577
Db 127 GCGGCAGCCAGGCTTATGCTAAGAGAAAGGTTTGACCCCTCTTCGACACCCACG 186

QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyAsnArgValLysAlaAla 597
Db 187 GATTTCAGGAATACCTGAGCAAAAGAGCAGAGATCTCTAACCGGCTGAAGAGCT 246

QY 598 LysLysValProLeuAspArgMetProTyAsnLeuGlnTyThrValGluValLysAsn 617
Db 247 AAGAGGTGCCACTTGATCGTATGCTTACATCTTCAGTATCTGAGAGTCAAAAC 306

QY 618 GlySerLeuIleIleProHisTyAspHisTyHisAsnIleLysPheGluTrpPheAsp 637
Db 307 GGTAGTTTAAATCATACCTCATTATGACCATTAACATCAATTTGAGTGGTTTGC 366

QY 638 GluGlyLeuTyThrGluAlaProLysGlyTyThrLeuGluAspLeuAlaThrValLys 657
Db 367 GAAGCCCTTTATGAGCCACCTAAAGGGPATAGTCTTGAGGATCTTTGGCGACTCTCA 426

QY 658 TyThrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677
Db 427 TACTATGTGAAA-----CCGGCG-----AACGCTAGT 453

QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697
Db 454 GACCATGTTCTGAAAATAAG-----GCAGACCAAGATAGTAAACCTGAT 498

QY 698 GluGluLysProGlnThrGluLysProGluGluThrProArgGluGluLysProGln 717
Db 499 GAAGATAAGCAACATGATGAAGTAAGTGAAGCAACTCACCCCTGAATCTGATCAAAAG 558

QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731
Db 559 AATCACGCTGTTTAAATCTTCACAGAGATAATCTTTATAAACAAGCACTGATACGGA 618

QY 732 GluSerProGluGluSerGluGluProGlnValGluThrGlu 745
Db 619 GAGACAGGAGGAAGCTGAAGATACCACAGATGAGGCTGAA 660

RESULT 36
US-10-324-143-6
; Sequence 6, Application US/10324143
; Publication No. US20030232976A1
; GENERAL INFORMATION:
; APPLICANT: HAMEL, JOSEPH
; APPLICANT: CHARLAND, NATHALIE
; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: MARTIN, DENIS
; APPLICANT: BLAIS, NORMAND
; APPLICANT: OUELLETTE, CATHERINE
; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 55190-56
; CURRENT APPLICATION NUMBER: US/10/324,143
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/341,252
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2721
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide
; OTHER INFORMATION: sequence
US-10-324-143-6

Alignment Scores:
Pred. No.: 7,076-65 Length: 2721
Score: 827.50 Matches: 168
Percent Similarity: 72.44% Conservative: 16
Best Local Similarity: 66.14% Mismatches: 29
Query Match: 20.59% Indels: 41
DB: 16 Gaps: 4

US-09-765-272A-66 (1-763) x US-10-324-143-6 (1-2721)

QY 498 GlnIleThrTyThrAspGluLeuGlnValAlaLysLeuAlaGlyLysTyThrThr 517
Db 4 CAAATTTACCTACCTGATGAGATTCAGGTAGCCAGTTGGCAGGCAAGTACACAA 63

QY 518 GluAspGlyTyThrPheAspProArgAspIleThrSerAspGluGlyAspAlaTyVal 537
Db 64 GAAGACGGTTATATCTTTGAT 84

QY 538 ThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArg 557
Db 85 -----CGTGGTTGGATTAAAGAGATAGTTTGTCTGAAGCTGAGAGA 126

QY 558 AlaAlaAlaGlnAlaTyThrAlaLysGluLysGlyLeuThrProSerThrAspHisGln 577
Db 127 GCGGCAGCCAGGCTTATGCTAAGAGAAAGGTTTGACCCCTCTTCGACAGACCCAG 186

QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyAsnArgValLysAlaAla 597
Db 187 GATTTCAGGAATACCTGAGCAAAAGAGCAGAGCTATCTAACCGGCTGAAGAGCT 246

QY 598 LysLysValProLeuAspArgMetProTyAsnLeuGlnTyThrValGluValLysAsn 617
Db 247 AAGAGGTGCCACTTGATCGTATGCTTACATCTTCAGTATCTAGAGTCAAAAC 306

QY 618 GlySerLeuIleIleProHisTyAspHisTyHisAsnIleLysPheGluTrpPheAsp 637
Db 307 GGTAGTTTAAATCATACCTCATTATGACCATTAACATCAATTTGAGTGGTTTGC 366

QY 638 GluGlyLeuTyThrGluAlaProLysGlyTyThrLeuGluAspLeuAlaThrValLys 657

Db 367 GAAGCGCTTATGAGCACCCTAAGGGTATAGTCTTGAGATCTTTGGGAGCTGTCAAG 426
 QY 658 TyrtYrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677
 Db 427 TACTATGTCGAGGG-----GGGAACGCTAGT 453
 QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697
 Db 454 GACCATGTCGTAATAAAG-----GCAGACCAAGATAGTAAACCTGAT 498
 QY 698 GluGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGln 717
 Db 499 GAAGATAGGAACATGATGAAGTAAAGTGAAGCCCAACTCACCTCGAATCTGATGATAAAGAG 558
 QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731
 Db 559 AATCAGCGTGGTTAAATCTTCAGCAGATATCTTTATAACCAAGCCTGATACGGAA 618
 QY 732 GluSerProGluSerGluGluProGlnValGluThrGlu 745
 Db 619 GAGACAGAGGAAGAGCTGAAGATACCAACAGATGAGGCTGAA 660

RESULT 37

JS-10-324-143-4
 ; Sequence 4, Application US/10324143
 ; Publication No. US20030232976A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HAMEL, JOSEE
 ; APPLICANT: CHARLAND, NATHALIE
 ; APPLICANT: BRODEUR, BERNARD R.
 ; APPLICANT: MARTIN, DENIS
 ; APPLICANT: BLAIS, NORMAND
 ; APPLICANT: OUELETTE, CATHERINE
 ; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
 ; FILE REFERENCE: 55190-56
 ; CURRENT APPLICATION NUMBER: US/10/324,143
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: 60/341,252
 ; PRIOR FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 160
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 816
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide
 ; OTHER INFORMATION: sequence
 JS-10-324-143-4

Alignment Scores:

Pred. No.: 3.88e-65 Length: 816
 Score: 822.50 Matches: 168
 Percent Similarity: 72.83% Conservatives: 17
 Best Local Similarity: 66.14% Mismatches: 28
 Query Match: 20.47% Indels: 41
 Gaps: 5

JS-09-765-272A-66 (1-763) x US-10-324-143-4 (1-816)

QY 498 GlnIleThrTyrtYrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrtYrThr 517
 Db 4 CAATTAACCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 63
 QY 518 GluAspGlyTyrtYrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrtVal 537
 Db 64 GAAGACGGTTATATCTTTCAT----- 84
 QY 538 ThrProHisMetThrHisSerHisTrpIleLysAspSerLeuSerGluAlaGluArg 557
 Db 85 -----ACTAGTTGGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGA 126

QY 558 AlaAlaAlaGlnAlaTyrtYrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577
 Db 127 GCGCAGCCAGCGGCTTATGCTAAAGAGAAAGGTTTGAACCCCTCTCTTCGACAGACACCAAG 186
 QY 578 AspSerGlyAsnThrGluAlaLysGlyValaGluAlaIleTyrtYrAsnArgValLysAlaAla 597
 Db 187 GATTCAAGAAATATCTGAGGCAAAAGGAGCAAGCTATCTACACCGCGGTGAAGCAGCT 246
 QY 598 LysLysValProLeuAspArgMetProTyrtYrAsnLeuGlnTyrtYrValGluValLysAsn 617
 Db 247 AAGAGGTGCCACTGATCGTATGCTTACATCTTCAGTATCTGTAGTAGTCAAAAC 306
 QY 618 GlySerLeuIleLeuProHisTyrtYrAspHisTyrtYrHisAsnIleLysPheGluTrpPheAsp 637
 Db 307 GGTAGTTTAAATCATACCTCATGACCATTAACCATCAATCAATTTGAGTGGTTTGAC 366
 QY 638 GluGlyLeuTyrtYrGluAlaProLysGlyTyrtYrLeuGluAspLeuAlaThrValLys 657
 Db 367 GAAGCGCTTATGAGGCACCTAAGGGGTATAGTCTTTGAGGATCTTTTGGCGACTGTCAAG 426
 QY 658 TyrtYrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677
 Db 427 TACTATGTCGAA-----CCGCGG-----AACGCTAGT 453
 QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697
 Db 454 GACCATGTCGTAATAAAG-----GCAGACCAAGATAGTAAACCTGAT 498
 QY 698 GluGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGln 717
 Db 499 GAAGATAAGGAACATGATGAAGTAAAGTGAAGCCCAACTCACCTCGAATCTGTGAAAAAGAG 558
 QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731
 Db 559 AATCAGCGTGGTTAAATCTTCAGCAGATATCTTTATAACCAAGCCTGATACGGAA 618
 QY 732 GluSerProGluGluSerGluGluProGlnValGluThrGlu 745
 Db 619 GAGACAGAGGAAGAGCTGAAGATACCAACAGATGAGGCTGAA 660

RESULT 38

US-10-324-143-5
 ; Sequence 5, Application US/10324143
 ; Publication No. US20030232976A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HAMEL, JOSEE
 ; APPLICANT: CHARLAND, NATHALIE
 ; APPLICANT: BRODEUR, BERNARD R.
 ; APPLICANT: MARTIN, DENIS
 ; APPLICANT: BLAIS, NORMAND
 ; APPLICANT: OUELETTE, CATHERINE
 ; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
 ; FILE REFERENCE: 55190-56
 ; CURRENT APPLICATION NUMBER: US/10/324,143
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: 60/341,252
 ; PRIOR FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 160
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 816
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide
 ; OTHER INFORMATION: sequence
 US-10-324-143-5

Alignment Scores:
 Pred. No.: 3.88e-65 Length: 816
 Score: 822.50 Matches: 168
 Percent Similarity: 72.83% Conservatives: 17

Best Local Similarity: 66.14% Mismatches: 28
 Query Match: 20.47% Indels: 41
 DB: 16 Gaps: 5

US-09-765-272A-66 (1-763) x US-10-324-143-5 (1-816)

```

QY 498 GlnIleThrTyrThrAspAspGluLeuGlnValAlaLysLeuAlaGlyTyrThrThr 517
D 4 CAATATTACCTACGATGATGAGATTGAGTTCAGGAGCAAGTTCGAGGCAAGTACACACA 63
QY 518 GluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVal 537
D 64 GAAGACGGTTATATCTTGAT----- 84
QY 538 ThrProHisMetThrHisSerHisTrpLeuLysAspSerLeuSerGluAlaGluArg 557
D 85 -----ACTAGTTGGATTAAATAAGATAGTTTCTCTGAAGCTGAGAGA 126
QY 558 AlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577
D 127 GCGGAGCCGAGGCTATGCTAAAGAGAAAGGTTTGAOCCCTCTTCGACAGACCACCAAG 186
QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAla 597
D 187 GATTGAGGAATATCTGAGGCAAAAGAGCAGAGCTATCTACACCGCTGAAGCAGCT 246
QY 598 LysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsn 617
D 247 AAGAAGGTGCCACTTGATCGTATGCTTACATCTTCAATCTTCACTACTGTAGAGTCAAAAAC 306
QY 618 GlySerIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAsp 637
D 307 GGTAGTTTAATCATACCTCATCATGACCATACCATAACATCAAAATTTGAGTGGTTTGAC 366
QY 638 GluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLys 657
D 367 GAAGCCCTTATGAGGACCTTAAGGGGTATAGTCTTGAGGATCTTTGGCGACTGTCAAG 426
QY 658 TyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677
D 427 TACTATGTCGAA-----COGCG-----AACGCTAGT 453
QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697
D 454 GACCATGTCGTAAATAAATAG-----GCAGACCAAGATAGTAAACCTGAT 498
QY 698 GluGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGln 717
D 499 GAAGATAGGACACATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 558
QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731
D 559 AATCAGCGTGTGTTTAAATCCTTCAGCAGATAATCTTTATAACCAAGCACTGATACGGAA 618
QY 732 GluSerProGluGluSerGluGluProGlnValGluThrGlu 745
D 619 GAGACAGGAGGAAGTGAAGATACCAACACAGATGAGCTGAA 660

```

RESULT 39

US-09-769-736-71

Sequence 71, Application US/09769736
 Publication No. US20030138775A1
 GENERAL INFORMATION:
 APPLICANT: Microbial Technics Limited
 APPLICANT: Le Page, Richard WF
 APPLICANT: Wells, Jeremy M
 APPLICANT: Hanniffy, Sean B
 TITLE OF INVENTION: Proteins
 FILE REFERENCE: PWC/P21089wo
 CURRENT APPLICATION NUMBER: US/09/769,736
 CURRENT FILING DATE: 2003-02-14
 PRIOR APPLICATION NUMBER: GB 9816335.5
 PRIOR FILING DATE: 1998-07-27

; PRIOR APPLICATION NUMBER: US 60/125163
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 71
 ; LENGTH: 1455
 ; TYPE: DNA
 ; ORGANISM: Streptococcus agalactiae
 US-09-769-736-71

Alignment Scores:

Pred. No.: 2 04e-39 Length: 1455
 Score: 542.50 Matches: 144
 Percent Similarity: 34.29% Conservative: 59
 Best Local Similarity: 24.32% Mismatches: 104
 Query Match: 13.50% Indels: 285
 DB: 13 Gaps: 13

US-09-765-272A-66 (1-763) x US-09-769-736-71 (1-1455)

```

QY 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
D 28 ACAGCAATCGGTGTTCTATTCCGACGGTAATCATTTCCACTTTATCACTATAAGGAT 87
QY 307 MetSerGluLeuGluLysArgIleAlaArgIleLeuProLeuArgTyrArgSerAsnHis 326
D 88 ATGCTCCATTAGAGTTAGAACCAACAAAGGATGGTG---GCAGAGCATAGAGGACATCAT 144
QY 327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346
D 144 ----- 144
QY 347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
D 145 -----ATTGAT----- 150
QY 367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
D 150 ----- 150
QY 387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
D 151 -----GCATTAGGGAAAAAAGATCTTACA 174
QY 407 AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSer 426
D 175 GAGAAACCAACGATATTTCTCATGAACCTTAATAAGGAA-----CCTCACACA 222
QY 427 AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeu 446
D 223 GAGGAAGAACACCAT----- 237
QY 447 AspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLys 466
D 237 ----- 237
QY 467 AspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIle 486
D 238 -----GCAGTAACA 246
QY 487 ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506
D 247 CGAAAGACCAACGTAAGCAAAACCAATAGCCAGATTGCTACAGATTGCTCAAGAAATT 306
QY 507 GlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProArg 526
D 307 GAAGAGGCAAAAAAGCTCGTAAATACACACATCTGATGGTTTACATTTTGTGCTAAA 366
QY 527 AspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr 546
D 367 GATATTAAAAAGATACAGGTATGATTCAGGTTTATGATTCATTCATATGACATGAGCATGG 426
QY 547 IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAla----- 564

```

```
db 427 GTACCAAGAAAGATTATCAGAGTCGGAATTAAGACAGCTCAAGAAATTTCTTTCAGGA 486
2y 564 -----
db 487 AATCTGAAGCAATCAAGACAAACCAAAACAGGTAAACAGCTCAAGAAATCTATGAG 546
2y 564 -----
db 547 GCAATTGAACCAAAAGCAATTTGTAACCTGAAGATTTATTATTGGAATTGCACAGCG 606
2y 564 -----
db 607 ACAGACTATAAGAAAGTACATTGTAATTCCTCATAAAGATCATTACCAATTATGTGAA 666
2y 565 -----
db 667 TTAATGTTTGTATGAAGAAAGATCTTTAGCTGATTCAGATAGACATATCTTTA 726
2y 569 -----
db 727 GAAGACTATTAGCTACGCTAAATATTATCATGATGCACCCAGAAAACGTCCTAAAGTT 786
2y 569 -----
db 787 GAAGATGGGGTAAAGATGCTGAAATTTATAAGAAAGGACTCTAATAAGCAGATAAA 846
2y 570 ---ThrProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGlu 588
db 847 CCAAGTCTGCACCACTGATTAATAATCAACATCAATCTAGTCACAAAACCTTAAGT 906
2y 589 Ala-----IleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMetPro 606
db 907 GCTGAGAAGTATTCAACAAAGCAACCAACCAAGAAAATTTGTACCGCTTGATATAAATGCT 966
2y 607 TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleleProHisTyrAsp 626
db 967 GCTCATGTCGATATGCAGTTTGGATTTGAAGATGATCAATTTGATTTCTCTCATCATGAT 1026
2y 627 HisTyrHisAsnIleLysPheGluTrpPheAspGlu---GlyLeuTyrGluAlaProLys 645
db 1027 CATATCATATATGTTCTTATGCTATGCTGTTTGAACAGGGTGGTTTATGGAAGCACCAGAA 1086
2y 646 GlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665
db 1087 GGTATACATTACCAACTCTTCTCAACAAATTAATACTACATGGAACATCTTAATGAA 1146
2y 666 ArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn 685
db 1147 TTACCA---AAAGAAAAGGGTTGGGACACACGACAGTGCAT-----AACAAAGGC 1194
2y 686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
db 1195 TCATAAAGACAAATAAAGCCAAAATATGCTCCAGATGAAGAACTGAA----- 1245
2y 706 ProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725
db 1246 -----
2y 726 ProThr-----GluGluProGlu 731
db 1258 GTAACCTCAACTATGTTTATGATGTTTAAAGGTTTCAGACCAAGAAAGACAGAA 1317
2y 732 GluSerProGluGluSerGluGluProGlnValGlu 743
db 1318 AAACAAGAAGATGAATCAGACTAGATGAATGAA 1353
2y
```

RESULT 40

JS-09-452-599-34/c

Sequence 34, Application US/09452599

Patent No. US20020055101A1

GENERAL INFORMATION:

APPLICANT: Bergeron, Michel G.

APPLICANT: Ouellette, Marc

APPLICANT: Roy, Paul H.

TITLE OF INVENTION: Specific and Universal Probes and Amplification Primers

TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial

TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro

FILE REFERENCE: 12287.31

CURRENT APPLICATION NUMBER: US/09/452,599

PRIOR FILING DATE: 1999-12-01

PRIOR APPLICATION NUMBER: 08/526,840

PRIOR FILING DATE: 1995-09-11

PRIOR APPLICATION NUMBER: 08/304,732

PRIOR FILING DATE: 1994-09-12

NUMBER OF SEQ ID NOS: 177

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 34

LENGTH: 841

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

US-09-452-599-34

Alignment Scores:

Pred. No.: 5,62e-37 Length: 841

Score: 512.00 Matches: 104

Percent Similarity: 93.10% Conservative: 4

Best Local Similarity: 89.66% Mismatches: 2

Query Match: 12.74% Indels: 7

DB: 9 Gaps: 2

US-09-765-272A-66 (1-763) x US-09-452-599-34 (1-841)

QY 653 LeuAlaThrValLysTyrValGluHisProAsnGluAlaGluProHisSerAspAsnGly 672

Db 839 TTGGCGACTGTCAAGTACTATGTGAAATCCAAACGACGTCG-CATTCAAGATATGTT 781

QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692

Db 780 TTTGGTAAACGCTAGCGACCATGTTCCAAAGAAACAAAATGGTCAAGCTGATACCAATCAA 721

QY 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluThrProArg 712

Db 720 ACGGAAAACCCAAAGGAGGAAACCTCAGACAGAAAACCTGAGGAGAAACCCCTCGA 661

QY 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu----- 729

Db 560 GAAGAGAAACCGCAAGCGAGAAACCGAGAGTCTCCAAACCAACAGAGAACCCAGAGAA 601

QY 730 -----ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysVal 747

Db 600 GAATCACCAGAGAAATCACCAGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAGTTTA 541

QY 748 GluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763

Db 540 ---AGAAAACCTGAGAGAGGCTGAAGATTTTACTTGGAAAAATCCAGAA 496

RESULT 41

US-10-121-120-34/c

Sequence 34, Application US/10121120

Publication No. US20030180733A1

GENERAL INFORMATION:

APPLICANT: Bergeron, Michel G.

APPLICANT: Ouellette, Marc

APPLICANT: Roy, Paul H.

TITLE OF INVENTION: Specific and Universal Probes and Amplification

TITLE OF INVENTION: Primers

TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial

TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro

FILE REFERENCE: 12287.31

CURRENT APPLICATION NUMBER: US/10/121,120

PRIOR FILING DATE: 2002-04-11

PRIOR APPLICATION NUMBER: 09/452,599

PRIOR FILING DATE: 1999-12-01

; PRIOR APPLICATION NUMBER: 08/304,732
 ; PRIOR FILING DATE: 1994-09-12
 ; NUMBER OF SEQ ID NOS: 177
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 34
 ; LENGTH: 841
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-10-121-120-34

Alignment Scores:
 Pred. No.: 5,62e-37 Length: 841
 Score: 512.00 Matches: 104
 Percent Similarity: 93.10% Conservative: 2
 Best Local Similarity: 99.66% Mismatches: 4
 Query Match: 12.74% Indels: 7
 DB: 15 Gaps: 2

US-09-765-272A-66 (1-763) x US-10-121-120-34 (1-841)

QY 653 LeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGly 672
 Db 839 TTGGCGACTGTCAAGTACTATGTCGAACATCCAAACGAAACGTCG-CATTCAGATATGCT 781
 QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
 Db 780 TTGGTAAAGCTAGCGACCATGTTCAAAAGAAACAAATAATGGTCAAGCTGATACCAATCAA 721
 QY 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluThrProArg 712
 Db 720 ACGGAAACCAACAGAGAGAAACCTCAGACAGAAACCTCAGAGAAACCCCTCGA 661
 QY 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu 729
 Db 660 GAACAGAAACCGCAAGCGAAACACAGAGTCTCCAAACCAACAGAGAGAAACCGAGAA 601
 QY 730 -----ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysVal 747
 Db 600 GAATCACAGAGAAATCACAGAGAAATCAGAGAAACCTCAGGTCGAGACTGAAAGGTTA 541
 QY 748 GluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763
 Db 540 ---AGAAACTGAGAGAGCTGAGATTACTTGGAAAAATCCAGAA 496

RESULT 42

US-10-104-047-536
 ; Sequence 536, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 536
 ; LENGTH: 3466
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-104-047-536

Alignment Scores:
 Pred. No.: 5,34e-07 Length: 3466
 Score: 190.50 Matches: 167
 Percent Similarity: 33.55% Conservative: 97
 Best Local Similarity: 21.22% Mismatches: 302
 Query Match: 4.74% Indels: 221
 DB: 16 Gaps: 40

US-09-765-272A-66 (1-763) x US-10-104-047-536 (1-3466)

QY 74 LysValProTyrAspAla-----IleIleSerGluGluLeuLeuMetLysAspProAsnTyr 92
 Db 1212 AAAGTGGCGGTAGATGTAATTAAGCAAGCAAGTGTGGAGAATAATCAAGCAATCTT 1271
 QY 93 GlnLeuLysAspSerAspIleValAsnGluLe-----LysGlyGlyTyrValle 109
 Db 1272 GCATCCAAAGAACAAAGAAACAATCAATGCTTCAACCATACCAAGGTTGGCAT----- 1325
 QY 110 LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg 129
 Db 1326 TCTGTATTTCAGCCATCAGTCTCTCTTGGTTGATCAAGATGGAACACCAAAATATATC 1385
 QY 130 ThrLysGluGluLysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAla 149
 Db 1386 ATCAACTACAGTCTTGAGCAGCCTAGCCAACTTCAAGTTGTTCTCCTCAAAATTTAAAAA 1445
 QY 150 AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyrIle 169
 Db 1446 GAAATCCAGTCCGTACAAACAGTTGTAAAGTGAAGATTACCAAGAGATCTTACTCTT 1505
 QY 170 PheAsnAlaSerAsp-----IleIle 176
 Db 1506 AAGTCTGAGAGAGCAAAAGCTTTGAAGGGGGGTGAATGATAGCAGTCTTCTTGTGT 1565
 QY 177 GluAsp---ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIlePro 195
 Db 1566 GATGATGTTCCAGAGATATTAATGCATCTCCAGATTAAAGCACTATGACCTA----- 1619
 QY 196 LysAsnGluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGln 215
 Db 1620 AAGCAGCTACTCAGCCTCTCCACTCCCTGAGCAGAGAGCT-----GAGAGCCTGAG 1673
 QY 216 GlySerArgProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 235
 Db 1674 TCTCTGTTTTCACGACTCTGGGAGTGGCAATTGTCCTCTAGTCAGCCACCTTTA--- 1730
 QY 236 GluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSer 255
 Db 1731 -----AAGAACCTCTTG 1742
 QY 256 SerLeuLeuArgGluLeuTyr-----AlaLysProLeuSerSerGluArgHisValGlu 272
 Db 1743 TCTCTCTAAAGCATATTATGCTTTGAATGCACAAACCAAGTCAGAGAGAG----- 1793
 QY 273 SerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaAlaGlyValAla 292
 Db 1794 -----CTCTCAAAATTTGCTGATTCAGTAAAC 1820
 QY 293 ValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLys 312
 Db 1821 CTACCA-----CTGGATGTAGTAAAAAG 1844
 QY 313 ArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerArg 332
 Db 1845 -----TGGTTTGAAGAGATGCAA 1862
 QY 333 ProGluGlnProSerProGlnSerThrProGluProSerProSerProGln----- 349
 Db 1863 GCTGGACAGATTTCAGTGCAGTCTTCTGAACCATCTTCTCTGAACCCAGGCAAGTAAAT 1922
 QY 350 ---ProAlaProAsn-----ProGlnProAlaProSerAsn---ProIleAspGlu 364
 Db 1923 ATCCCTGCCAAGAACATGATCAGCTCAATCTGCAAAATGCAATGACCCAGGAGCAGC 1982
 QY 365 -----LysLeuValLysGluAlaValArgLysValGlyAsp 376
 Db 1983 ACAGTAAATCTACAAAGTCTTTTGAAGATGACTAACTCCCGAGTTTACCAGTGGGATCA 2042
 QY 377 GlyTyrValPheGluGluAsnGlyValSerArgTyrIleProAla---LysAspLeuSer 395
 Db 2043 -----ACCACCAATGGTTTCCAGAGTAGTACACCATCCCATCACCTCTAAAC 2090

```
396 AlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLys 415
      ::::: |||
2091 CTTTCCTCATCCAGAAATACACAGGGTTACTGTGTACACAGCTGAGGGTGCAACAAGAG 2150
      ::::: |||
416 LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLys 433
      ::::: |||
2151 CCACAAGTAGAACCTCTTGATCTTTTCACTACCAAGCAACACGAGGAGATATTAGAAAG 2210
      ::::: |||
434 AlaTyrAspLeuLeuAlaArgIleGlnAspLeuLeuAspAsnLysGlyArgGlnVal 453
      ::::: |||
2211 TCA-----ACTATCACTAGTTGTACCAG-----AACAGTGTATTATCTGTCTGC 2252
      ::::: |||
454 AspPheGluAlaLeuAspAsnLeuGluArgLeuLysAspVal**SerAspLysVal 473
      ::::: |||
2253 CAGGAAGAACCCCTTG--AACTGTCTTGGCGAAAAAGAGGCCACAAAAGGAC----- 2303
      ::::: |||
474 LysLeuVal***Asp-----IleLeuAlaPheLeuAlaProIleArgHisProGlu 490
      ::::: |||
2304 AGTTGTGTTACAGACTCAGAACCCAGTTGTAATGTAATCCCAACCAAGTGCACACCCATA 2363
      ::::: |||
491 ArgLeuGlyLysProAsnAla-----GlnIleThrTyrThrAspAspGlu 505
      ::::: |||
2364 AATATCGGTATACCTACAGTCAGTCCGCCAGTTACCCACATCTGTGGCCATCTGTCACGAC 2423
      ::::: |||
506 IleGlnVal-----AlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyr 521
      ::::: |||
2424 AACAGTGTTCATGCTTAAGAGCGCTAGCTGCCAATAAGCAACG----- 2468
      ::::: |||
522 IlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMet 541
      ::::: |||
2469 -----ATTCTGACTCCCCAGGTG 2486
      ::::: |||
542 ThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln 561
      ::::: |||
2487 GCATACACCTAC-----TCACTACGCTCAGCGCCCTGCAGTCCAA 2525
      ::::: |||
562 AlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsn 581
      ::::: |||
2526 -----GAACACCCCTTGAAAGTGATCCAGCCAAATGGAAT 2561
      ::::: |||
582 -----ThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLys 595
      ::::: |||
2562 CAGGATGAAGACAAAGATACCTAGCTCAGAGGAGTATCAATGTAGAGGATCAGATGAC 2621
      ::::: |||
596 AlaAlaLysLysValProLeuAspArgMet-----ProTyrAsnLeuGlnTyrThrVal 613
      ::::: |||
2622 TCTGATTCACCGCCCAAAAGAAATGCGGAGACAGAAATGGAATGTATGCTGT 2681
      ::::: |||
614 GluVal-----LysAsnGlySerLeuIleIleProHisTyrAspHis 627
      ::::: |||
2682 GATTTGTGACAGATATTCAAAAGAGTAGTTCATTATTGAGACATAAATATGACAC 2741
      ::::: |||
628 -----TyrHis 629
      ::::: |||
2742 ACAGGTAAGACCTCATGAGTGTGGAATCTGTAAGGCGATTTAAACACAAACATCAT 2801
      ::::: |||
630 AsnIleLysPheGluTrpPheAspGluLysLeuTyrGluAlaProLysGlyTyrThrLeu 649
      ::::: |||
2802 TTGATTGAA-----CACATGCGATTACATCTCGGAGAAAGCCCTATCAATGT 2849
      ::::: |||
650 GluAspLeuLeuAlaThrValLysTyr-----TyrValGluHisProAsn----- 664
      ::::: |||
2850 GACAAATGTGAAAGCGCTTCTCACACTCTGGGTCTTATTCTCAACACATGATGATCATGCG 2909
      ::::: |||
665 -----GluArgProHisSerAspAsnGlyPheGlyAsn 675
      ::::: |||
2910 TACTCTACTGTAAAGAGACCGGAAGACGTGACAGACACAGACGAGGAGGAGCAGG 2969
      ::::: |||
676 -----AlaSerAspHisVal---GlnArgAsnLysAsnGlyGlnAlaAspThr 690
      ::::: |||
2970 CCTGAATTCCTCTCGAATGAGCAGTGGGTGCGAGGGGTCTCCCTCAGCGGCGACTCG 3029
      ::::: |||
691 AsnGlnThrGluLys---ProSerGluLysProGlnThrGluLysProGluGlu 709
```

```
Db 3030 GACGAGAGAGAGAGTTTGCACAGGAGAGAGATGACAGACAGTGAAGAGAGAGAGGAG 3089
      ::::: |||
QY 710 ThrProArgGluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu 729
      ::::: |||
Db 3090 GAGGATTAAGAGATGAGAAATTCGAGGAAGAAATGTGAAAGAACCAACCAAGGGGAT 3149
      ::::: |||
QY 730 ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysValGluGlu 749
      ::::: |||
Db 3150 GAGGAAGAGGGGAGGAGGAGGAGGAGAA-----GTGGAGAGAGAGAGGTAGAGAG 3200
      ::::: |||
QY 750 LysLeuArgGluAlaGluAsp 756
      ::::: |||
Db 3201 GCAGAGAATGAGGAGAGAGAA 3221
      ::::: |||

RESULT 43
US-10-087-192-233
; Sequence 233, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-233

Alignment Scores:
Pred. No.: 1,25e-06 Length: 3494
Score: 186.50 Matches: 166
Percent Similarity: 33.55% Conservative: 98
Best Local Similarity: 21.09% Mismatches: 302
Query Match: 4.64% Indels: 221
DB: 13 Gaps: 40

US-09-765-272A-66 (1-763) x US-10-087-192-233 (1-3494)
QY 74 LysValProTyrAspAla-----IleIleSerGluGluLeuMetLysAspProAsnTyr 92
      ::::: |||
Db 1252 AAAGTCGGGTAGATGATTAATGTAAGCAAGTGTGAGATAATCAAGCCAACTTT 1311
      ::::: |||
QY 93 GlnLeuLysAspSerAspIleValAsnGluIle-----LysGlyGlyTyrValIle 109
      ::::: |||
Db 1312 GCATCAAGAGACAAACAATCAATGCTTACCCTACACAGGTGGCCAT----- 1365
      ::::: |||
QY 110 LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg 129
      ::::: |||
Db 1366 TCTGTTATTTCAGCCATCAGTCTCTCTTGTGATCAAGATGGAACCAACCAAAATATC 1425
      ::::: |||
QY 130 ThrLysGluGluIleLysArgGlnLysGlnLysSerHisAsnHisAsnSerArgAla 149
      ::::: |||
Db 1426 ATCACTACAGTCTTGAGCAGCCCTAGCACTTCAAGTTGTTCTCTCAAAATTTAAAAAA 1485
      ::::: |||
QY 150 AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyrIle 169
      ::::: |||
Db 1486 GAAATCCAGTCGTACAAACAGTGTGTAAGAGTGAAGTACAGAGATCTTACTGTT 1545
      ::::: |||
QY 170 PheAsnAlaSerAsp-----Ileile 176
      ::::: |||
Db 1546 AAGTCTGAGAGACAAAGCTTTGAAGGGGGGTGNAATGATAGCACTTGTCTCTGTGT 1605
      ::::: |||
```


TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS

FILE REFERENCE: 11613.56US11

CURRENT APPLICATION NUMBER: US/10/159,563

CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: US 10/133,937

PRIOR FILING DATE: 2002-04-25

NUMBER OF SEQ ID NOS: 444

SOFTWARE: PatentIn version 3.1

SEQ ID NO 295

LENGTH: 5316

TYPE: DNA

ORGANISM: Homo sapiens

US-10-159-563-295

Alignment Scores:

Pred. No.: 2,21e-06 Length: 5316

Score: 186.50 Matches: 166

Percent Similarity: 33.55% Conservative: 98

Best Local Similarity: 21.09% Mismatches: 302

Query Match: 4.64% Indels: 221

DB: 16 Gaps: 40

US-09-765-272A-66 (1-763) x US-10-159-563-295 (1-5316)

```

QY 74 LysValProTyrAspAla---IleLeuSerGluLeuLeuMetLysAspProAsnTyr 92
DB 1252 MAAGTGGCGGTAGATGCTAATGTAATAGCGAAGTGTGGAGAATAATCAAGCAATCTT 1311
QY 93 GlnLeuLysAspSerAspIleValAsnGluLeu-----LysGlyGlyTyrValle 109
DB 1312 GCATCCAAAGAACCAAGAACCAATCAATGCTTCCACCATCAACCAAGTGGCCAT- 1365
QY 110 LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg 129
DB 1366 TCTGTTATTTCAGCCATCAGTCTTCTCTGTTGATCAAGATCGAACCAACCAATATATC 1425
QY 130 ThrLysGluGluIleLysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAla 149
DB 1426 ATCACTACAGTCTTGAGCGCTAGCCACTCAAGTTGTTCTCCTCAAAATTTAAAAA 1485
QY 150 AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIle 169
DB 1486 GAAATCCAGTCCGTACAAACAGTTGTAAGAGTGAAGTACAGAGATCTTACTGTT 1545
QY 170 PheAsnAlaSerAsp-----IleIle 176
DB 1546 AAGTCTGAGAGACAAAAGCTTTGAAGGGGGGTGAATGATAGCACTTCTCTCTGTGT 1605
QY 177 GluAsp---ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIlePro 195
DB 1606 GATGATGTCCAGGAGATATTAAATGCACTTCCAGATTAAGCACTATGACCTA----- 1659
QY 196 LysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGln 215
DB 1660 AAGCAGCCTACTCAGCCTCTCCACTCCCTCGCAGCAGAGCT-----GAGAAGCCTGAG 1713
QY 216 GlySerArgProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSer 235
DB 1714 TCTCTGTTTTCATCAGCTAGCTGGAGATGGCAATTTGTTCTCTAGTCAGCCACCTTTA- 1770
QY 236 GluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSer 255
DB 1771 -----AAGACCTCTTG 1782
QY 256 SerLeuLeuArgGluLeuTyr-----AlaLysProLeuSerGluArgHisValGlu 272
DB 1783 TCTCTCTAAAGCATATATGCTTTGAATGACACCAACCAAGTCAGAGAG----- 1833
QY 273 SerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAla 292
DB 1834 -----CTCTCAAAATTTGCTGATTTCAGTAAAC 1860
QY 293 ValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLys 312

```

```

DB 1861 CTACCA-----CTGATAGTAAAAAAG 1884
QY 313 ArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTyrValProAspSerArg 332
DB 1885 -----TGGTTTGAAGAAGTCAA 1902
QY 333 ProGluGlnProSerProGlnSerThrProGluProSerProSerProGln----- 349
DB 1903 GCTGGACAGATTTCAGTGCAGTCTTCTGAACCATCTTCTCTCTGAACAGGCAAAAGTAAAT 1962
QY 350 ---ProAlaProAsn-----ProGlnProAlaProSerAsn---ProIleAspGlu 364
DB 1963 ATCCCTGCCAAGAACAAATGATCAGCCTCAATCTGCAAAATGCAAAATGAACCCGAGGACAGC 2022
QY 365 -----LysLeuValLysGluAlaValArgLysValGlyAsp 376
DB 2023 ACAGTAATCTCAAAAGTCTTTGAGATGATACTCCCAAGTTTACCAAGTGGATCA 2082
QY 377 GlyTyrValPheGluGluAsnGlyValSerArgTyrIleProAla---LysAspLeuSer 395
DB 2083 -----ACCACCAATGGTTCAGAGAGTAGTACACCATCCCATCATCCTCTAAAC 2130
QY 396 AlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLys 415
DB 2131 CTTTCTCATCCAGAAATACACAGGGTTTACTTGTACACAGCTGAGGGTGCACAAAGAGAG 2190
QY 416 LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLys 433
DB 2191 CCACAAAGTAGAACCTCTTGTACTTCTACCAAGCAACAGGAGAAATATTAGAAAGG 2250
QY 434 AlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnVal 453
DB 2251 TCA-----ACTATCAGTCTGTTTACCAG-----AACAGTCTTATTCTGTCTC 2292
QY 454 AspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysVal 473
DB 2293 CAGGAAGAACCCCTTG---AACTTGTCTTTCGCAAAAAGGAGGCCACAAAGAGAC----- 2343
QY 474 LysLeuVal***Asp-----IleLeuAlaPheLeuAlaProIleArgHisProGlu 490
DB 2344 AGTTGTGTACAGACTCAGAACCACTGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 2403
QY 491 ArgLeuGlyLysProAsnAla-----GlnIleThrTyrThrAspAspGlu 505
DB 2404 AATATCGTATACCTACAGTCACTGCCAGTTCACCAATCGTGGCCATTGCTGACCAG 2463
QY 506 IleGlnVal-----AlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyr 521
DB 2464 AACAGTGTTCATGCTTAAGAGCGCTAGTGCATAAAGCAACG----- 2508
QY 522 IlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMet 541
DB 2509 -----ATTCTGATCCCCAGGTG 2526
QY 542 ThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln 561
DB 2527 GCATACACCTAC-----TCACTACGCTCAGCCTCGAGTCCAA 2565
QY 562 AlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsn 581
DB 2566 -----GAACCACTCTGAAAGTATCCAGCAAAATGGAAT 2601
QY 582 -----ThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLys 595
DB 2602 CAGGATGAAGACAAAGATACCTAGCTCAGAGAGTATCAAAATGTAGAGGATCAAGATGAC 2661
QY 596 AlaAlaLysLysValProLeuAspArgMet-----ProTyrAsnLeuGlnTyrThrVal 613
DB 2662 TCTGATTCTACCGCCCAAAAGAAATGCGGAGACAGAAAAATGGAATGATGCTGTTGT 2721
QY 614 GluVal-----LysAsnGlySerLeuIleIleProHisTyrAspHis 627

```


[illegible]

Qy	710	ThrProArgGluGlnLysProGlnSerGluLeuPheProGluSerProLysProThrGluGlu	729	
Db	3146	GAGGATAAAGAGATGGAGAATTTCACAGCAAGAAAAAGAAATGTGAATAAACACCAAGGGGAT	3205	
Qy	730	ProGluGluSerProGluGlnSerGluGluProGlnValGluThrGluLysValGluGlu	749	
Db	3206	GAGGAA-----GAGGAGGAGGAGGAGGAAGAAGTGGAAAGAGAGGTTAGAAGAG	3256	
Qy	750	LysLeuArgGluAlaGluAsp	756	
Db	3257	GCAGAGAATGAGGGAGAGAA	3277	
RESULT 46				
US-10-232-561-9				
; Sequence 9, Application US/10232561				
; Publication No. US20030119772A1				
; GENERAL INFORMATION:				
; APPLICANT: Genetta, Thomas				
; TITLE OF INVENTION: Methods and compositions useful for				
; FILE REFERENCE: CHOP 00-99				
; CURRENT APPLICATION NUMBER: US/10/232,561				
; CURRENT FILING DATE: 2002-08-30				
; PRIOR APPLICATION NUMBER: 60/317,300				
; PRIOR FILING DATE: 2001-09-05				
; NUMBER OF SEQ ID NOS: 12				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 9				
; LENGTH: 5126				
; TYPE: DNA				
; ORGANISM: Human				
US-10-232-561-9				
Alignment Scores:				
Pred. No.:	7.35e-06	Length:	5126	
Score:	33.04%	Matches:	167	
Percent Similarity:	33.04%	Conservative:	96	
Best Local Similarity:	20.98%	Mismatches:	290	
Query Match:	4.49%	Indels:	243	
DB:	15	Gaps:	40	
US-09-765-272A-66 (1-763) x US-10-232-561-9 (1-5126)				
Qy	74	LysValProTyrrAspAla---	lleIleSerGluGluLeuLeuMetLysAspProAsnTyrr	92
Db	1351	AAAGTGCGGTAGATGGTAATGTAATAAGCAAGTCTTGAGAAATAATCAAGCCAATCTT	1410	
Qy	93	GlnLeuLysAspSerAspleValAsnGluIle-----	LysGlyGlyTyrrValIle	109
Db	1411	GCATCCAAAAGACAAGAAACAATCAATAGTCTCACCCCATACAAACAGGTGCCAT----	1464	
Qy	110	LysValAsnGlyLystrTyrrValTyrrLeuLysAspAlaAlaHisalaAspAsnIleArg	129	
Db	1465	TCTGTTATTTCAGCCATCAGTCTTCCTTTGGTTGATCAAGATGGAACAACCAAATATATC	1524	
Qy	130	ThrLysGluGluIleLysArgGlnLysGlnUargSerHisAsnHisAsnSerArgAla	149	
Db	1525	ATCAACTACAGTCTTGAGCGACCTAGCCCACTCAAGTTGTTCTCTCAAAATTTAAAAAAA	1584	
Qy	150	AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrrThrTrpAspAspGlyTyrrIle	169	
Db	1585	GAAATCCAGTCGCTCAACACAGTTGTAAAGTAGTAAAGTACCAGAGATCTTACTGTT	1644	
Qy	170	PheAsnAlaSerAsp-----	lleIle	176
Db	1645	AAGTCTGAGAGGACAAAAGCTTTGAAGGGGGGTGAATGATGACACTTGCTCTCTGTGT	1704	
Qy	177	GluAsp---ThrGlyAspAlaTyrrIleValProHisGlyAspHistyrHisTyrrIlePro	195	
Db	1705	GATGATTGTCGAGAGATATTATATGCACTTCCAGATTAAAGCACTATGACCTA-----	1758	
Qy	196	LysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrrTrpAsnGlyLysGln	215	

Alignment Scores:	7.35e-06	Length:	5226
Pred. No.:	180.50	Matches:	167
Score:	33.04%	Conservative:	96
Percent Similarity:	30.98%	Mismatches:	290
Best Local Similarity:	20.98%	Indels:	243
Query Match:	4.49%	Gaps:	40
DB:	15		

US-09-765-272A-66 (1-763) x US-10-232-561-9 (1-5126)

Qy	74	LysValProTyrAspAla---	llelleSerGluGluLeuLeuMetLysAspProAsnTyr	92
Db	1351	AAAGTCGGGTAGATGTTAATGAAGCAAGTCTTGAGAAATTAATCAAGCAATCTT	1410	
Qy	93	GlnLeuLysAspSerAspIleValAsnGluIle-----	LysGlyGlyTyrValIle	109
Db	1411	GCATCCAAAGAACCAAGAAACAATCAATGCTTCACCATACAAAGGTGCCAT-----	1464	
Qy	110	LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg	129	
Db	1465	TCGTGTTATTTCACCATCAGTCTCTCTTGTTGATCAGATGGAACCAACCAAAATTATC	1524	
Qy	130	ThrLysGluGluIleLysArgGlnLysGlnUArgSerHisAsnHisAsnSerArgAla	149	
Db	1525	ATCAACTACAGTCTGTGACGAGCGCTAGCAACTTCAAGTTGTTCTCCTCAAAATTATAAAAAA	1584	
Qy	150	AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIle	169	
Db	1585	GAAATATCCAGTCCCTACAAACAGTTGTAAAGAGTGAAGTAGTACCAGAAGATCTTACTGTT	1644	
Qy	170	PheAsnAlaSerAsp-----	llelle	176
Db	1645	AAGTCTGAGAAGACAAAAGCTTTGAGGGGGGTGAATGATAGCACTGTCTCTGTGT	1704	
Qy	177	GluAsp---ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIlePro	195	
Db	1705	GATGATTGTCGAGAGATATTAAATGCATCTCCAGAATTAAAGCACTATGACCTA-----	1758	
Qy	196	LysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGln	215	

1759	AGCAGCCTACTCAGCCTCTCCACTCCCTGCGAGCAGAAGCT-----GAGAAAGCCTGAG	18181
216	GlySerArgProSerSerSerSerSerSerTyAsnAlaAsnProAlaGlnProArgLeuSer	235
1813	TCCTCTGTTTCATCAGCTACTCGAGATGGCAATTTGCTCTCTAGTCAGCCACTTTA---	1869
236	GluAsnHisAsnLeuThrValThrProThrTyHisGlnAsnGlnGlyGluAsnIleSer	255
1870	-----AAGAACTCTCTTG	1881
256	SerLeuLeuArgGluLeuTyr-----AlaLysProLeuSerGlnuArgHisValGlu	272
1882	TCCTCTCTAAAGCAATATATGCTTTGAATGCAACCAAGTCAGCAAGAG-----	1932
273	SerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAla	292
1933	-----CTCTCAAAATTTGCTGATTGAGTAAC	1959
293	ValProHisGlyAsnHisTyHisPheIleProTyrGluGlnMetSerGluLeuGluLys	312
1960	CTACCA-----CTGGATGTAGTAAAAAAG	1983
313	ArgIleAlaArgIleIleProLeuArgTyArgSerAsnHisTrpValProAspSerArg	332
1984	-----TGGTTTGAAGAAGTCAA	2001
333	ProGluGlnProSerProGlnSerThrProGluProSerProSerProGln-----	349
2002	GCTGGACAGATTTCAGTGCAGTCTTCTGAACCATCTCTCTGAAACGAGCAAGTAAT	2061
350	ProAlaProAsn-----ProGlnProAlaProSerAsn-----ProIleAspGlu	364
2062	ATCCCTGCCAAGAACAATGATCAGCCTCAATCTGCAAAATGCAAAATGAACCCAGGACAGC	2121
365	-----LysLeuValLysGluAlaValArgLysValGlyAsp	376
2122	ACAGTAAATCTACAAAGTCCTTTGAAAGATGATCACTCCCGCTTTTACAGTGGGATCA	2181
377	GlyTyrValPheGluGluAsnGlyValSerArgTyIleProAla-----LysAspLeuSer	395
2182	-----ACCACCAATGGTCCAGAAAGTAGTACACATCCCATCCCTCTTAAC	2229
396	AlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGlnSerLeuSerHisLys	415
2230	CTTTCTCATCCAGAAATACACAGGGTTACTTGTACACAGCTCAGGGTGCACAAGAAGAG	2289
416	LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLys	433
2290	CCAAAGTAGAAGCCTCTGTGATCTTCTACTACCAAAAGCAACAGGAGAA-----	2337
434	AlaTyAspLeuLeuAlaArgIle-----HisGlnAspLeuLeuAspAsnLysGlyArg	451
2338	-----TTATTAGAAAGAAATTCCTTTACACAGACAGTGTATTCTGTCCAGGAAGA	2388
452	GlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAsp	471
2389	ACCCTTGAA-----CTTGTCTTGGCAAAAAAGGACCCAAAAAGGAC	2430
472	LysValLysLeuVal***Asp-----IleLeuAlaPheLeuAlaProIleArgHis	488
2431	-----AGTTGTGTATACAGACTCAGAACCAAGTTGTTGTAATGTAATCCCAAGTCGCAAC	2484
489	ProGluArgLeuGlyLysProAsnAla-----GlnIleThrTyrThrAsp	503
2485	CCCATAAATATCGTATACCTATACAGTCACGTGCCAGTTACCCAAATCTGGCCCATGCT	2544
504	AspGluIleGlnVal-----AlaLysLeuAlaGlyLysTyThrThrThrGluAsp	519
2545	GACCAGAACAAGTGTTCATGCTTAAGAGCGCTAGCTGCCAATAAGCAAAACG-----	2595
520	GlyTyIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyValThrPro	539
2596	-----ATTCTGATTCCC	2607

Qy	540	HisMetThrHisSerHisTrpIleLeuLysAspSerLeuSerGluAlaGluAAla	559
Db	2608	CAGGTGGCATACACCTAC-----TCAACTACGGTCAGCCCTGCA	2646
Qy	560	AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThrAspHisGlnAspSer	579
Db	2647	GTCCAA-----GAACCAACCTTTGAAGTGTATCCAGCCAAAT	2682
Qy	580	GlyAsn-----ThrGluAlaLysGlyValaGluAlaIleTyrAsnArg	593
Db	2683	GGAAATCAGGATGAAAGACACAGATACTAGCTCAGAAGGAGTATCAAATGTAGAGGATCAG	2742
Qy	594	ValLysAlaAlaLysLysValProLeuLysAspArgMet-----ProTyrAsnLeuGlnTyr	611
Db	2743	AATGACTCTTGAATCTTACACCGCCCAAAAGAAATCGGAAGACAGAAAAATGGAATGTAT	2802
Qy	612	ThrValGlnVal-----LysAsnGlySerLeuIleIleProHisTyr	625
Db	2803	GCTTGTGATTTGTGTGACAAAGATATTCCAAAGAGTAGTTCATTTATGTAGACATAAATAT	2862
Qy	626	AspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLys	645
Db	2863	GAACACACACAGGTAAAGACCTCATGAGTGT-----GGAATC-----TCTAAATAAG	2907
Qy	646	GlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGlu	665
Db	2908	GCATTTAAACACAAACATCATTTG-----ATTGAACACATGCGATTACATCTTGAGAA	2961
Qy	666	ArgPro-----HisSerAspAsnGlyPheGly	674
Db	2962	AAGCCCTATCAATGTGACAAATGTGGAAGCGCTTCTCACACTCT-----GGG	3009
Qy	675	AsnAlaSerAspHisValGln-----	681
Db	3010	TCTAGTTCTCAACACATGAATCATCGCTACTCCATCTGTAAGAGAGAAAGCGAAGAACGT	3069
Qy	681	-----	681
Db	3070	GACAGCACAGACAGGAGGAGCGGCGCTGAAATCTCTCGAATGAGCACGTGGGTGCC	3129
Qy	682	ArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLys---ProSerGluGluLys	700
Db	3130	AGGGCGTCTCCCTCACAGGGCGACTCGGACGAGAGAGAGATTGTGCAAGGGAAGAGGAT	3189
Qy	701	ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys	720
Db	3190	GAAGACACAGTGAATAAAGGGAAGGAGGAGGAGGATTAAGAGATGGAAGAAATTCGAGGAGAA	3249
Qy	721	ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro	740
Db	3250	AAAGAATGTGAATAAACACAGGGATGAGGAA-----GAGGAGGAGGAGGAGAA	3300
Qy	741	GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGlnAsp	756
Db	3301	GAAGTGGAGAAGACAGGTTAGAAGAGGACAGGAATCAGCGAGAAGAA	3348

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 45846

LENGTH: 3570

TYPE: DNA

ORGANISM: Saccharomyces cerevisiae

S-10-369-493-45846

alignment Scores:

red. No.: 5,53e-06 Length: 3570

core: 179.50 Matches: 180

percent Similarity: 33.08% Conservative: 126

est Local Similarity: 19.46% Mismatches: 351

very Match: 4.47% Indels: 269

B: 16 Gaps: 40

S-09-765-272A-66 (1-763) x US-10-369-493-45846 (1-3570)

```
Y 10 AlaGlyGlnValLysLysGluSer-----AsnArgValSerTyr 22
b 787 GCAGGTAATAAATAAATCAAGCATCTCGAGGAAGATATCGTGTCCAGATCAGCAACC 846
Y 23 IleAspGlyAspGlnAlaGlyGlnLys-----AlaGluAsn 34
b 847 GTGGATCCAGATAAAACAAAGCAGCAAGCTCTATTGTCGGAACTCCTATATATGTCCACC 906
Y 35 LeuThrProAspGluValSerLysArgGluGlyIleAsnAlaGluGln***ValIleLys 54
b 907 ATGACTCCAAACAGAGTTACT----- 927
Y 55 IleThrAspGlnGlyTyrValThrSerHis----- 64
b 928 CTTACCGGTGAGAGTGCTGTTCTGAAACGACACGACATAAAACGAATTTGTTCTCTGCT 987
Y 65 ---GlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIleIleSerGlu 83
b 988 TCAGCTTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1041
Y 84 GluLeuLeuMetLysAspPro-----AsnTyrGlnLeuLys----- 95
b 1042 GACATTAATAGAAACACCCCGTACCACTCCAGATAAATCAATCAGGAATGTTGAATATC 1101
Y 96 -----AspSerAspIleValAsnGluIleLysGlyGly----- 106
b 1102 CCCGAGCCAGAAATCTCAGAGATTGAAAGAGATTAAAGAGGGACCTTCCAGTCGGGCC 1161
Y 107 -----TyrValIleLysVal-----AsnGly 113
b 1162 TCCATTTTACCGCTAAAGCTGCAAGATTCCCAATGAAACCACTTACTTTGGAATATGTC 1221
Y 114 LysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGlu 133
b 1222 GAGAACTAC---GAATCCGACGATAGCTCAAGTAGCAGAAACCAAGAACTCTCGAAACA 1278
Y 134 IleLysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaVal 153
b 1279 GAACACATTCARAGGCTTCCTTCGACAGTAGTCAAGCTCCATCGCAGATAAT----- 1332
Y 154 AlaAlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSer 173
b 1333 -----AATGGCTCT 1341
Y 174 AspIleIleGluAsp---ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHis 192
b 1342 CCGGTGAAATAATAGTCTCTCGTGATGCC-----ATGCCACAAATGTGCAATTAGCA 1395
Y 193 TyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsn 212
b 1396 GAACCTCCCAAGCGAAGCAATACCTCGATCACCATACTTCAAAATGGCGAATCCTGGGA 1455
Y 213 GlyLysGlnGlySerArgProSerSerSerSerTyr----- 225
b 1456 AAGCAACAGAACATCAACACCAAGAAAGATTCTCGGAAACTATAGTGGAAAGAAA 1515
```

```
QY 226 ---AsnAlaAsnProAla----- 230
Db 1516 TCTCAGGCGGAACCTTCCGGTATTGTGGAACCGAAAGAAATGACTAAATTTCTAGATGAT 1575
QY 231 ---GlnProArgLeuSerGluAsnHisAsn-----LeuThrValThrPro 244
Db 1576 AATCAGGTAAAGAGAAAGAAAGATACGAATGATAAACTTTTAGAAAAGAAATTTGCGCG 1635
QY 245 ThrTyrHisGlnAsn-----GlnGlyGluAsnIleSerSerLeuLeu 258
Db 1636 ACAATTCACATATGATCAGCAATACTTGCATCATCTGATAGTCTAATGGGACGTTG 1695
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHis----- 270
Db 1696 AAATCACTAGTCGGAAAGTTAGTTCAACCAACACGCTTCAAGGAAGATGGCACTATC 1755
QY 271 -----ValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286
Db 1756 ATTAACGGCACCATTTGAGGACGATGGCAATGATGATGAATGATGATGATGATGATGAT 1815
QY 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
Db 1816 -----ATTGTTCTCTCAAGATTCAGATTCCTCTAGTTTCCCAAGTCAGAT 1860
QY 307 MetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHis 326
Db 1861 CTTTTCAAATGATCGAAGTGATGATACAGATCTACGCATGCTGTTTAAGGGCAAA--- 1917
QY 327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346
Db 1918 -----AATTCGAGAACTTCGGGTAAATTCCAAAAATTCAAAGCCTTACCAACGCTT 1968
QY 347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
Db 1969 TTGATAAAGACATAGACATTTCTAAGCCAGATTCCAAGAAATATTTTGTCCCGCAGCAAC 2028
QY 367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
Db 2029 CMAGAAGTCGGCTAAGAGG----- 2049
QY 387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
Db 2050 -----GCTGCCCAATTTATGGCAGGTGCGCAAGAAAAACGAA 2085
QY 407 AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuPro----- 424
Db 2086 GTCCCAACAGAAAAGTACTGAAGATTCTAGTTCCGCGCAGCTCCACAGAGATGAGTCAGAT 2145
QY 425 -----SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeu 437
Db 2146 AGTGGGATTGAAACAGATTTCTCATCAGATGATGATTTTAAAGGAAAAATATGTCGGTG 2205
QY 438 LeuAlaArgIleHisGlnAspLeu----- 445
Db 2206 CCTAATAATGGGCCCCAAGACATTTCTTACATTTCTTGAAGGCGAGTGTGTTCCAGTG 2265
QY 446 LeuAspAsnLys-----GlyArgGlnValAspPheGluAlaLeuAspAsnLeuGlu 463
Db 2266 AAAGACTCAAAAGATAATCAATAAAGAAAGTTGATGAAGAAAGAAACGCAAAAAGAGATAGT 2325
QY 464 ArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeu 483
Db 2326 CAAAAAAGTTCACCGCTGAGTGAATCTTCAGTAACAAATAGCAAAATAGTGAAACAATG 2385
QY 484 AlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAsp 503
Db 2386 GCGAAAAGCTTTTATCCGAATTCAAACGAAGCAAAATGAGCT-----ACAAG 2436
QY 504 AspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePhe 523
Db 2437 GTGGAAACTAAGCCAGCAACACAGCATCGTCTTTCTGTTGTTGGGG----- 2487
```

524 AspProArgAspIleThr-----SerAspGluGlyAspAlaTyrVal 537
Db |||||
2488 TCCTCTCTGTTGCTACTAAAGGAACACCTCTTCAATGAAGAGAAAT----- 2538
QY |||||
538 ThrProHisMetThrHisSerHisTyrPileLysLysAspSerLeuSerGluAlaGluArg 557
Db |||||
2539 -----CGTAAATGTTAAACGGAAGCCCAAGAT 2568
QY |||||
558 AlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577
Db |||||
2569 GAGTCTCGCAGATGTAGTCAGCAGCAAGGAACAACTCTAGGTTGTCAGATTGAAA 2628
QY |||||
578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaLysTyrAsnArgValLysAlaAla 597
Db |||||
2629 ---AGCCCAACATA-----GGAGGTGAAGACTTGAACAAGAACGACGAGGGTCC 2676
QY |||||
598 LysLysValPro-LeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAs 617
Db |||||
2677 AAGGAA---CCAGAAAAGCATCTGCCATATATACAGGATGCTAACGATAAAATAACTCG 2733
QY |||||
617 nGlySerLeuIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAs 637
Db |||||
2734 AAGGAAAAGAAAGACTCCAAAGCAACACAGGTACT-CAGAAAAAGTTAAATACACAGA 2792
QY |||||
637 p-----GluGlyLeuTyrGluAlaProLys-----GlyTyrThrLeuGluAs 651
Db |||||
2793 TCCTTAAAGAGGCAATGTTACAGCTTCTTAAAGCCAGTCCAAATGATAAATTGAAGGA 2852
QY |||||
651 pLeuLeuAla-----ThrValLysTyrTyrValGluHisPro-----AsnG 665
Db |||||
2853 TTTCAGGCTAAATTTACTAATAGTAAACACACTCGTGCACCTGGAATAAATTTCTAATGA 2912
QY |||||
665 uArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAs 685
Db |||||
2913 AAAA-----AATAATCTAGTCACACAGATGATCTTCTAGCTCCTCAGCTTCTTC 2963
QY |||||
685 nGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
Db |||||
2964 CACAGAAAGAGAGTCTAGTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3023
QY |||||
705 sProGluGlu-----GluThrProArgGluGluLysProGlnSerGluLysPr 721
Db |||||
3024 GAAGGTAGNAGAGTGTGTTAAATACCAAGAGAGAACCCAGTTAGATCTTCTCATCATAGAT 3083
QY |||||
721 oGluSerProLysProThr-----GluGluProGluGluSerPro-- 734
Db |||||
3084 TGAAGCCCTCTCTCCATCTGTTTAAACAAAATCAATGCCACCCCTGATAAAATCCAGT 3143
QY |||||
735 -----GluGluSerGluGluProGlnValGluThrGlu----- 745
Db |||||
3144 TACACAGCTAATGGACATGTCGTCTCTCTCGTCAGTTAAATCCAAAGACCACCTCAAAATCC 3203
QY |||||
746 -----LysValGluGluLysLeuArgGluAlaGluAs 756
Db |||||
3204 GTCTAGTATACCTACGACCTTCCAGAAAAGTTGCTCTCTTCTTGTAGTTCATTAATCTGA 3263
QY |||||
756 pLeuLeuGlyLys 760
Db |||||
3264 TCTTGTTTCAAGA 3276

RESULT 48

US-09-842-552-98
; Sequence 98, Application US/09842552
; Patent No. US20020055628A1
; GENERAL INFORMATION:
; APPLICANT: The Regents of The University of California
; TITLE OF INVENTION: MULTICLOCUS REPETITIVE DNA SEQUENCES FOR GENOTYPING BACILLUS ANTHR
; FILE REFERENCE: S-89,687
; CURRENT APPLICATION NUMBER: US/09/842,552
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/199,911
; PRIOR FILING DATE: 2000-04-26

; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 98
; LENGTH: 2522
; TYPE: DNA
; ORGANISM: Bacillus anthracis
US-09-842-552-98

Alignment Scores:

Pred. No.: 4, 7e-06 Length: 2522
Score: 178.00 Matches: 119
Percent Similarity: 34.06% Conservative: 52
Best Local Similarity: 23.71% Mismatches: 151
Query Match: 4.43% Indels: 180
DB: 9 Gaps: 27

US-09-765-272a-66 (1-763) x US-09-842-552-98 (1-2522)

QY 382 GluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGly 401
Db |||||
12 GAAATGCTCAATTAAATTTATTGAAAAAAGATAGTGAATCTGGTCAACTTCTACCAAGGT 71
QY 402 Ile-----AspSerLysLeuAla----- 407
Db |||||
72 GCAAAATTTGATGTTATCGATAAAGATGGAAAGTTGTGAAACAATTTGTTACAGATGAT 131
QY 408 LysGlnGluSerLeuSerHisLysLeuGlyAla-----LysLysThrAsp 422
Db |||||
132 AAAGTGAAGCTTTATCGAAACAACCTCCAGTTGGAAGCTATACATTAAAGAAAGTAGAA 191
QY 423 LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeu--LeuAlaArgIle 441
Db |||||
192 GCACCG-----AAAGATATGAATTAATCACTAGTTCAGTT 227
QY 442 HisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPhe----- 455
Db |||||
228 TCTGTTGATGATGAGGCTAAATAAAGTAGTACTGTAGTGTGGTGAATAAAAGATCCCC 287
QY 456 GluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal**SerAspLysValLysLeu 475
Db |||||
288 GAAAGTAACAGGTCAATTTGAAGTGGTGAAGTAGATGCAAAATGATAAAACGAAATTG 347
QY 476 Val***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysPro 495
Db |||||
348 TTATCA----- 353
QY 496 AsnAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyr 515
Db |||||
354 GGTGCAGAAATTCGAAGTGTATAAAGATGCAAAAAGGTAGCAGAACTGAAA----- 404
QY 516 ThrThrGluAspGlyTyrIlePheAspProArg-----AspIleThrSerAsp 531
Db |||||
405 ACAGTGAGATGGAAAGATGATGTCACCGAAATTTACCGCTAGGTGAATACACAGTGA 464
QY 532 GluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr----- 546
Db |||||
465 GAAACGAAGCACCGCGGCTACAGCTTTCAGATAAAGATGGAAGTAACAATTCAA 524
QY 546 ----- 546
Db 525 AACGAGAAAGAGTAGTAAAGTAGAGCGCAGAAAACGAAAAAATCTTAGTTTCTCTACAA 584
QY 547 ---IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyr----- 563
Db |||||
585 ATTATTAATGATGATGATAAAGATCAACAGAAACGCTTAGCAGCGCGCAGAAATTACATTG 644
QY 564 -----AlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577
Db |||||
645 AAAGATGTAAAGCAATGTTGTAAGAAAGAGAAATTACA-----ACAGAT----- 689
QY 578 AspSerGlyAsnThrGluAlaLysGly-----AlaGluAlaIleTyrAsnArgVal 594
Db |||||
690 AAGTCTGGAACCTGTTAAAGTAGACGAGCTTGTGCGGGTGAATATACCTTAGAAGAAACA 749

QY 595 LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614
Db 750 AAGAGCCGCAAGAGTTATAAGCA-----TTAAGAGTAAACATCGAA 791
QY 615 ValLys----- 616
Db 792 GTAACAGTAGTAGCAACAGAGTAGTAAACAGACGCTGTGTAATGAAAGAGTGAAGAA 851
QY 617 -----AsnGlySerLeuIleProHisTyrAspHisTyrHisAsnIleLys----- 632
Db 852 GAAATTACAGGCAATTAGAAATTACAAAGTAGATGCTAATGATATAAATAAATAA 911
QY 633 -----PheGlu-----Tyr----- 635
Db 912 GCAGGCGCAGTGTGAAATTTGGAAGAGCGCAACAAATCGATACATTAACATCAGAT 971
QY 636 -----PheAspGluGlyLeuTyr----- 641
Db 972 GAAATGGTAAAGCAATTTGGAAGAGCGCAACAAATCGATACATTAACATCAGAT 1031
QY 642 GluAlaProLysGlyTyrThrLeuGluAsp-----LeuLeuAlaThrVal 656
Db 1032 CAAGCCCGAGAGGTATGAGTTATCTGTATAGGAATCGAATTTCGATTCTTAATCAA 1091
QY 657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAla 676
Db 1092 AATTTGAAGTTGTAAGACTTCAAAATTACAAATAAAGAAAGCAAGAGTCCAGAG 1151
QY 677 SerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysPro 696
Db 1152 AATCCAGGCGCAAGAAACAGAAAGAGCCAGGTGAA-----GAAACAGAAAGCCG 1199
QY 697 SerGlu-----GluLysPro-----GlnThrGluLysProGluGluThrProArg 712
Db 1200 GGTGAAGAAACAGAAAGAGCCAGGCGAAGAAACAGAAAGCCAGGTGAAGAAACA 1253
QY 713 GluGluLysPro-----GlnSerGluLysPro-----GluSerProLysProThrGlu 728
Db 1254 ---GAAAGCCAGGTGAAGAAACAGAAAGCCAGGCGAAGAAACAGAAAGCCAGGTGAA 1310
QY 729 GluProGluLysSerProGluGluSerGluGluProGlnValGluThrGluLysValGlu 748
Db 1311 GAAACAGAAAGCCAGGTGAAGAAACAGAAAGAGCCAGGCGAAGAAACAGAAAGCCAGGT 1370
QY 749 GluLys 750
Db 1371 GAAGAA 1376

RESULT 49
US-09-816-669A-13
; Sequence 13, Application US/09816669A
; Patent No. US20020137019A1
; GENERAL INFORMATION:
; APPLICANT: GARABEDIAN, Michael
; APPLICANT: TANEJA, Samir
; APPLICANT: HITTELMAN, Adam
; APPLICANT: MARKUS, Steven
; TITLE OF INVENTION: METHOD FOR SCREENING TRANSCRIPTIONAL COREGULATORY PROTEINS OF
; TITLE OF INVENTION: TRANSCRIPTION FACTORS, AND ANDROGEN RECEPTOR TRANSCRIPTIONAL CO
; TITLE OF INVENTION: PROTEINS AS TARGETS FOR ANDROGEN RECEPTOR-DEPENDENT DISEASES
; FILE REFERENCE: GARABEDIAN-1, 1A
; CURRENT APPLICATION NUMBER: US/09/816, 669A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/225,618
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/191,768
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 8588
; TYPE: DNA

ORGANISM: Human
US-09-816-669A-13
Alignment Scores:
Pred. No.: 2,78e-05 Length: 8588
Score: 177.50 Matches: 138
Percent Similarity: 30.01% Conservative: 70
Best Local Similarity: 19.91% Mismatches: 201
Query Match: 4.42% Indels: 285
DB: 9 Gaps: 30
US-09-765-272A-66 (1-763) x US-09-816-669A-13 (1-8588)
QY 216 GlySerArgProSer-----SerSerSerSerTyrAsnAlaAsn 228
Db 3163 GGTCTGTAGCCTTCCATGCTCCTCCACGCATGCTTCAGATGCCAGAGGACGCCACC 3222
QY 229 ProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrHisGln 248
Db 3223 AAGGCC-----CTGCTGGAGAACTTTGGCTTTGAGTTGGTTCATCCAGTATATAG 3273
QY 249 Asn-----GlnGlyGluAsnIleSerSer 256
Db 3274 AACAGCAGAGGTGCAGAAAAAGAAATGGAAAGACTGACCCAGGAGAGAACCTGGAAAG 3333
QY 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276
Db 3334 CTC-----GAGTGTGACTCCTCGCGCAAG 3357
QY 277 IlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGly 296
Db 3358 TTGTTTTCACACATCTTGATTTTAAAG-----AGTCATCAA 3393
QY 287 AsnHisTyrHis-----PheIleProTyrGluGlnMetSerGlnLeuGluLysArgIle 314
Db 3394 GAGCAGCTTTCATCAGATTTACTTTCTTCAACACAGCTCGAGAGGTTTGCACAAACATAC 3453
QY 315 Ala-----ArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAsp 330
Db 3454 AGAGACCACTAGATAAACTGTACCCTGAGGCGCCAGACC-----CCAGAG 3501
QY 331 SerArgProGluGlnProSerProGlnSerThrProGluProSerProSerProGlnPro 350
Db 3502 CCACCACCACTCCCTCCCTCCACCTCCACCTCCACCTCCCGCAGCGCGCTCAGCG 3561
QY 351 AlaProAsnProGln-----ProAlaProSerAsnProIleAspGluLysLeuValLysGlu 369
Db 3562 GCGTCCACACCAAGCCATCCCGCATCAGCCCAACCATCCTCCTCCTCCTCCTCCTCCTC 3621
QY 370 AlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSerArgTyrIle 389
Db 3622 GCCCAGCATCAGTGCC-GCTCACCAGCTCTCCATGCCGATGGAGCTGCCCATCTCTC 3680
QY 390 ProAlaLysAspLeu-----SerAlaGluThrAlaAla 400
Db 3681 GCCGCTGATGTCAGACGATGCCGTCGACACCTTGGCGGCTCAGCTACCTCCCGCAGCT 3740
QY 401 GlyIle----- 402
Db 3741 GGGACCTGTGGAGCCTCTGCTTGGGAGCCTGGCCAACTCTACCAAGCATCAGCTCAATCC 3800
QY 402 ----- 402
Db 3801 AACCTGTCTCAGCAGCAGAAACAGAGGCTTCGCCACAGGATCAGATGATCAGCTCG 3860
QY 402 ----- 402
Db 3861 AGCTTGTGGCAATATTTTGTACATTAACTCCCTCCAGTGAAGCAATAAAGAGAT 3920
QY 403 -----AspSerLysLeuAlaLysGlnGluSerLeu 412
Db 3921 GGCAGACAAAGTCGGGTGTCCTCCCAAGAAAGTGTATCAGCACTGTTTCAGGAACACTCTCT 3980

3454 AGAGACCCTACGATAAAGCTGTAACCACTGAGGCCCCAGACC-----CCAGAG 3501
331 SerArgProGluGlnProSerProGlnSerThrProGluProSerProSerProGlnPro 350
3502 CCACCACCACCTCCCCCTCCACCCCTCCACCCCTCCACCCCTCCACCCCTCCAGCG 3561
351 AlaProLenProGln---ProAlaProSerAsnProIleAspGluLysLeuValLysGlu 369
3562 GCGTCACACACGAGCCATCCCGCATACGCCCCACCCATCACCCTCACCAATTGCACCG 3621
370 AlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSerArgTyrIle 389
3622 GCCCAGCCATCAGTGCC-GCTCACCCAGCTCTCCATGCCGATGAGTGCCCATCTTC 3680
390 ProAlaLysAspLeu-----SerAlaGluThrAlaAla 400
3681 GCCGCTGATGATCGACGATGCCGCTGCAGACCTTCGCGCTCAGCTACCCCGCAGCT 3740
401 GlyIle----- 402
3741 GGGACCTGGAGGCTTGCTGCTGGGACCTGGCCCAACTCTACCAGCATCAGCTCAATCC 3800
402 ----- 402
3801 AACCTGTCTCCAGCAGCAGAAACAAGAGCCCTGCACCCAGGATCAGATGATCAGCTCCG 3860
402 ----- 402
3861 AGCTCTGGCGCAATATTTTGACATTAACTACACTCCCCCAGTCGACAGCAATAAAGAGAT 3920
403 -----AspSerLysIleAlaLysGlnGluSerLeu 412
3921 GGCAGACAAGTCGCGGTGCCCCAGAAAGTGATCAAGCACTGGTTCCAGAACACTCTCTT 3980
413 Ser-----HisLys 415
3981 CAAGAGAGGCAGCGGTACAAGGACTCCCTTACAATTCAAGTTCAGTAATCCTCTCATCACCG 4040
416 LeuGlyAlaLysLysThrAspLeuProSer----- 425
4041 CCTGGAGGAGCTCAAGATTGACTCCCGGCCCTTCCCGGCACTCCAAAGCAGGAGTA 4100
426 -----SerAspArgGluPheTyrAsnLysAlaTyrAspLeuAla-ArgIleHisGly 443
4101 CTGGGGAACGAAGAGCTCTCAAGACAAGGTTTACGACTTACCAGCTGAGGGTCTTACA 4160
443 nAspLeuLeuAspAsnLysGly-----ArgGlnValAspPheGluAlaLeuAspAsnLe 461
4161 GGACTTCTTCGATGCCAATGCTTACCACCAAGGATGATGAATTTGAGCAACTCTCTAATTT 4220
461 uLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal**AspIleLeuAl 481
4221 ACTGAACCTTCCACCCGAGTG-----ATAGTGGT 4250
481 aPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTy 501
4251 GTGGTTTCAGAATGCCGACAGAGCCAGCAAGCAATATCAGAAT----- 4296
501 rThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTy 521
4297 -----CAGGGAGAGGCGAAA-----GATGGA- 4317
521 rIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMe 541
4318 -----GAGCGGGTGGCTTACAAATGAT----- 4341
541 tThrHisSerHisrIleLysLysAspSerLysGluAlaGluArgAlaAlaAlaGly 561
4342 -----AGATACATTCGAACGAACCTGAATCACCAGTGCAAAAATAGTAGCT 4391
561 nAlaTyrAlaLys-----GluLysGlyLeuThrProProSerThr 574
4392 GGTGTTTTCAGCGGCATCTTTGATCTCATCAAGCACCAAGCAAGCTGGTTTACAGGATGA 4451

Qy	574	rAspHisGln-----AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaLeuTyr	591
Db	4452	GGATGAGAGGGCGAGGACGACGCAAAATGAGGATTCATGGATGCCATGGAAATCCT	4511
Qy	591	rAsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsn-----	608
Db	4512	GAGCGCTACAGCTCATCTCGAGTACCCCGATGCCCTCACAGCTTACAGCCGCCGAGC	4571
Qy	609	-----LeuGlnTyrThrValGluValLysAs	617
Db	4572	ACCATCAGGCCAATAATACAGCTTCTCCGCTTCTTGACGCTTACAGCGAGCGTGGAGGA	4631
Qy	617	nGlySerLeuLeuLeuProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAs	637
Db	4632	ACTGGCCACCTTC-----AATTCAAAAACAGCAGCGCGGA	4667
Qy	637	pGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLys	657
Db	4668	T-----GAGAAACCAAG-----	4680
Qy	657	sTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSe	677
Db	4681	-----CTGGCGGAGCTCCAGTGCACAGCAACCAAA-----	4713
Qy	677	rAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSe	697
Db	4714	-----ACCCAGAAAAAGCAAGACAA-----CCAAA	4739
Qy	697	rGluGluLysProGlnThrGluLysProGluGluGlu-----ThrProArgGluGluLys	715
Db	4740	GCCAGAGCTGCAGCAGCAGAGCAGCCGAGAGAGACCAACACITCCCCAGCAGAGCT	4799
Qy	715	sProGln-----SerGluLysProGluSerProLysPro--	726
Db	4800	CCCCAGCTGGTGTCCCTGCTGTGTCACAGCCTCTCCACAGAGGCCCTCCACA	4859
Qy	727	-ThrGluGluProGluGluSerProGluGluSerGlu	738
Db	4860	GTGCCCCCTTACCCAGTCGAGCCCGCCAGTCTCTCCAG	4896

Search completed: May 1, 2004, 08:41:51
Job time : 1035 secs